The Dissertation Committee for Sugandha Dhar Tuladhar certifies that this is the approved version of the following Dissertation:

Confidence Intervals for Computable General Equilibrium Models

Committee:

Peter J. Wilcoxen, Supervisor

Don Fullerton

David A. Kendrick

Leon S. Lasdon

Daniel T. Slesnick

Confidence Intervals for Computable General Equilibrium Models

by

Sugandha Dhar Tuladhar, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2003

Acknowledgements

I wish to acknowledge the assistance and encouragement I have received from many indispensable individuals during the course of my studies. Without their help and support I could not have completed my graduate studies.

I wish to thank The Fulbright Commission for supporting me for my graduate studies in Operations Research and Industrial Engineering at The University of Texas at Austin. I sincerely appreciate the intellectual guidance of the faculty members of the Operations Research and Industrial Engineering Program.

Dr. Peter J. Wilcoxen, who graciously agreed to be my supervisor, must be singled out for special thanks. It was his research interest and idea that led me to the topic and the culmination of this study. I appreciate his patience when I made no sense; encouragement when I reached an impasse; and funding when I needed assistantship. I owe a debt of gratitude for all this and the amount of time he devoted in guiding me through. Above all, it is his genuine desire to help that I appreciate the most, thank you Dr. Wilcoxen.

I was fortunate enough to take Dr. David Kendrick's computational economics class, as an engineering student, that led to where I am now. It would have been impossible to come this far without his valuable advice. I am grateful for his suggestions to improve my dissertation work.

To the other members of my dissertation committee: Dr. Don Fullerton, Dr. Daniel Slesnick, and Dr. Leon Lasdon. I am also truly grateful for valuable insights, suggestions, and steady encouragement.

Thanks are also due to the local Nepalese community for making our family feel homely during our stay in Austin. In particular, Kamal and Meena Adhikary for accepting us as one of their family members. I would also like to thank Kathie Julian of the Asian Development Bank and Professor John Adams, University of Virginia for providing continuous support and encouragement to complete the thesis while on assignment for the Bank in Nepal. Most importantly, I would like to thank my wife, Anju, and my two kids (Syona and Ronit) for their patience to make this study a reality.

Finally, financial support for the dissertation from U.S. Environmental Protection Agency under Cooperative Agreement CR824259 and the National Science Foundation under Grant SBR-9321010 is gratefully acknowledged.

S.D.T.

Confidence Intervals for Computable General Equilibrium Models

Publication No._____

Sugandha Dhar Tuladhar, Ph.D.

The University of Texas at Austin, 2003

Supervisor: Peter J. Wilcoxen

Computable general equilibrium (CGE) models have expanded from being a simple theoretical tool to a widely accepted policy evaluation tool. Despite recognizing that model parameters involve uncertainty, virtually all modelers report their results without confidence intervals. This obscures the uncertainty inherent in the models and gives the impression that the results are far more certain than they actually are.

CGE models with calibrated parameters and econometric CGE models using only the mean value of the parameters share a common flaw: their results are point estimates only, with no indication of the range of possible variation. A better analysis would include confidence intervals that communicate the underlying uncertainty. This would allow the policy makers to understand the precision of the results.

In this dissertation a tractable formal technique for calculating confidence intervals is presented. The results from this approach are compared with sensitivity analysis, an alternative method sometimes used for assessing uncertainty. It is shown for the models presented that sensitivity analysis can produce misleading results and that the confidence intervals are feasible to compute and qualitatively superior.

Next, the technique is applied to an econometric intertemporal general equilibrium model of the US economy to examine a current policy issue. The strong form of the double dividend hypothesis, which asserts that revenue-neutral substitution of an environmental tax for a distortionary income tax can improve welfare, is tested. The intertemporal equivalent variation (EV) for the policy is calculated. Unlike other studies, however, the 95 percent confidence interval for the EV is presented. The mean EV is slightly negative but the confidence interval is large and includes zero, so the model neither supports nor rejects the double dividend hypothesis. In addition, the short-run and the long-run intratemporal EV is calculated and compared to the intertemporal EV. The result implies that the long-run result supports the double dividend hypothesis even though the short-run does not. Finally, I present a detailed analysis of the general equilibrium effects that yield these distinct and contradictory results.

In sum, this dissertation provides an econometric view of CGE modeling and statistical testing of CGE results that is acceptable to econometricians. It attempts to answer criticisms of CGE modeling and the wider challenge to empiricism in economics (Whalley, 1985).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Computable general equilibrium (CGE) models that are used for policy evaluation are complex and can involve thousands of equations and hundreds of parameters. The policy results calculated by such models depend, often heavily, on the numerical values of the parameters. However, none of the parameters are known precisely. At best, they are econometrically estimated from historical data; at worst, they are chosen by calibrating the model to reproduce a single data point. Either way, CGE results depend on imperfect estimates of unknown parameters. As a result, they should be viewed and interpreted as econometric estimates of the effects of the policies being considered. A crucial question, therefore, is how precise these estimates are.

To illustrate how the precision of a CGE result depends on the precision of the underlying parameter values, consider a typical production function. Suppose, for example, that output, Q, is assumed to be a Cobb-Douglas function of three inputs, intermediate goods (X), capital (K), and labor (L) with constant returns to scale:

$$Q = AX^{\alpha}K^{\beta}L^{1-\alpha-\beta} \tag{1.1}$$

The function involves three unknown parameters: A, α, and β, and a natural way to quantify them would be to estimate them econometrically using equation 1.1.

The econometric specification of the equation would usually be written in log form with an additive stochastic term, ε:

$$q = a + \alpha x + \beta k + (1 - \alpha - \beta) l + \varepsilon \tag{1.2}$$

Variables q, x, k, and l are the logarithms of Q, X, K and L respectively. The output, q, typically reported by a CGE model would be given by equation 1.2 evaluated using the estimated parameter values: $\hat{a}$, $\hat{\alpha}$, and $\hat{\beta}$, and the mean disturbance term ε. Since the mean value of ε is zero, the log of output can then be written as:

$$\hat{q} = \hat{a} + \hat{\alpha} x + \hat{\beta} k + \left(1 - \hat{\alpha} - \hat{\beta}\right) l \tag{1.3}$$

Although CGE studies generally ignore it, $\hat{q}$ has a variance as well. The variance will depend on the variance-covariance matrix of the parameter estimates, and on the residual variance in the disturbance term. Both components are likely to be non-negligible in a typical CGE model. The parameters are likely to have large variances if the number of observations in the sample is small and the explanatory variables are closely related. Similarly, if the model's equations fit poorly, then the residual variance will be large. Large variances in the parameters and the error term will lead to a large variance for $\hat{q}$, or imprecision in the model's results. A natural and effective way to communicate the underlying precision of the results would be to report a confidence interval for $\hat{q}$. However this practice

is rarely adopted in CGE modeling: virtually all CGE results are presented without confidence intervals.

On one hand, the lack of confidence intervals is partly due to the way many CGE models are constructed. Most authors parameterize their models by selecting a few important parameters from a literature search and then choosing the remainder by calibrating the model to a benchmark data set (Mansur, et al., 1983, Shoven and Whalley, 1992). In some situations, such as modeling a developing country, calibration of some of a model's parameters may be unavoidable due to inadequate data. The result of calibration, however, is a set of parameter values that are essentially based on a single data point. There is no variance-covariance matrix for the parameters, and it is thus impossible to compute confidence intervals for the parameters or for the model's results.

On the other hand, some CGE models do exist in which parameters have been determined by estimation. However, even these models generate results without confidence intervals. The reason, in part, is that it has generally been considered impractical to compute confidence intervals for large models. Building on Smith's (1993) approach, Kim (2003) models uncertainty by integrating a dynamic CGE model into an optimal control framework. The hybrid model is solved using established method of passive-learning stochastic optimal control theory.[1]

[1] Other modeling techniques, for example, control theory base economic models (Kendrick, 1981) successfully incorporated uncertainty. Kim (2003) showed the importance of relative degree of associated parameter uncertainty in a model by integrating a traditional CGE model within a control theory framework. Harrison and Vinod, (1989) propose a method for estimating the population mean of the distribution of solution values from applied general equilibrium model subject to uncertainty using randomized factorial sampling designs. Harrison, et al., (1993), suggests conditional systematic sensitivity analysis and unconditional systematic sensitivity

Thus, virtually all current CGE models share a common flaw: they report results that are point estimates only and fail to report the corresponding confidence intervals. They lack, in other words, any systematic measure of the precision of their results.

At best, CGE studies report sensitivity analysis, which shows the effects of changes (often arbitrary) in a few selected parameters on key endogenous variables. The analysis is assumed to provide a reasonable approximation to the true precision of the model's results. However, ad hoc selection of the parameters for sensitivity analysis cannot ensure that they are, in fact, the key parameters of the model. Furthermore, the choice of parameters for sensitivity analysis could bias the result toward a particular outcome (Pagan, et al., 1985). A much better approach would be to compute confidence intervals instead.

Confidence intervals would allow policy makers to understand the precision of a model's results in a familiar and intuitive way. For an illustration, consider two hypothetical policy-relevant statements:

Statement 1:

> *"The revenue from the tax will be 700 million dollars."*

Statement 2:

> *"The revenue is most likely to be 700 million dollars and with 95 percent probability it will be between 660 and 740 million dollars."*

---

analysis to see how robust applied general equilibrium analysis are. Non-parametric approach, Monte Carlo simulation, is widely used for analyzing the effects of uncertainty. Channing (1996) introduces another method, using systematic sensitivity analysis via gaussian quadrature, to accomplish a similar objective. Webster, et al., 1998 calculates probability distributions of key climate change variables due to uncertainties in climate sensitivity using Deterministic Equivalent Modeling Method.

The second statement is clearly a more realistic, credible, and intellectually honest result. It acknowledges the fact that uncertainty exists regarding both parameter values and the choice of equations used in the model, and it gives a corresponding range of possible variation in the results.

In addition, confidence intervals would be a valuable tool in comparing results from alternative models. Different models often produce substantially different results when used to examine a given policy. For example, Paul L. Joskow cites a wide range of results in his comments on Jorgenson, Slesnick, and Wilcoxen (1992):

> *The resulting carbon tax required to meet the CO_2 is $16 per ton carbon, in 1989 dollars, which is associated with a loss of GNP of 0.5 percent in JSW model. For a similar constraint case applied to the Manne-Richels model produces a carbon tax of $300 and a loss in GNP of 3.2 percent. Most of the other models that have analyzed similar CO_2 constraints yield tax rates and GNP losses between these two extremes.*

In the absence of confidence intervals, policy makers frequently react to these differences by deciding that one of the models must be right and the others must be wrong. Much effort and debate is then devoted to choosing which model is right. However, the differences between models could very well be within the confidence intervals of the results. In that case, choosing one set of results and discarding the rest is worse than useless: there is no need to do it and it throws away vitally important information about the underlying uncertainty in the estimated effect of the policy. Thus, presenting confidence intervals would significantly improve the level of debate in comparing alternative models.

The goal of this dissertation is to demonstrate the importance of parameter and residual uncertainty in CGE modeling and to present a method that can be

routinely performed to calculate confidence intervals in CGE models. The basic approach is straightforward. First, the model is linearized using the Johansen method. Next, the linearized model is used to calculate the variance-covariance matrix of the model's endogenous results based on the variance-covariance matrices for the parameter estimates and residuals. The method is applied to a series of increasingly complicated and more realistic models. In addition, this approach is also applied to a real policy issue: the double dividend hypothesis. The rest of the dissertation is outlined as follows.

Chapter 2 introduces the notion of a confidence interval with a simple model. The model contains a simple linear equation with two parameters. The confidence interval is compared with the results from a typical sensitivity analysis and differences between the two are investigated. This chapter also illustrates sensitivity analysis in a graphical form.

Chapter 3 develops a small, static general equilibrium model. The parameters are econometrically estimated using aggregate data for the United States. A tractable formal approach for calculating confidence intervals is elaborated in detail and the resulting confidence intervals are compared with the outcome of a Monte-Carlo simulation. Again, comparisons with sensitivity analysis results indicate why confidence intervals are qualitatively superior.

Chapter 4 outlines a more realistic econometric intertemporal computable general equilibrium of the United States. The model contains five agents: three industries, one consumer, and a government. It has twenty-three behavioral parameters and eighteen residuals. Confidence intervals accounting for

uncertainty in the parameter estimates and for unexplained or residual uncertainty in the estimating equations are calculated and analyzed for both the steady state and the full intertemporal base case solution.

Chapter 5 extends the analysis to confidence intervals for policy experiments. In particular, the approach is applied to a double dividend hypothesis and used to conduct a formal statistical test of the strong form of the hypothesis. A variety of interesting results are presented and explored further via additional simulations.

Chapter 6 presents the dissertation's conclusion. It argues that routine calculation and reporting of confidence intervals for CGE results is valuable, technical feasible, and long overdue. In short, just as it is unacceptable for an econometric study to report estimated coefficients without standard errors, it should now be unacceptable for CGE results to be reported without confidence intervals.

Chapter 2: An Econometric Perspective on Uncertainty and Sensitivity Analysis

2.1 INTRODUCTION

Econometric forecasting models and CGE models are the two basic streams approaches that are widely used for policy evaluation. Forecasting models are often nonstructural and data-intensive. The regressors in the equations of forecasting models are sometimes included subjectively, thus often yielding extraordinarily good fit, as indicated by R^2 values close to one.[2] A typical equation from a forecasting model, in this case the DRI Model, is shown below. The example illustrates extraordinary goodness of fit for the chosen explanatory variables. Standard errors are given in parentheses and the equation's adjusted R^2 is 0.9913.

lnexpcs = - 2.5 + 0.72 lnpic + 1.26 lndinc
(0.17) (.0.05) (0.28)

- 0.71 log(0.6 pcncs/pc + 0.4 pcncs(-1)/pc(-1_)) + 0.06 log(W(-1)/pc)
(0.03) (0.02)

The dependent variable, *lnexpcs*, is the log of real consumption of clothing and shoes per capita. The explanatory variables are defined as follows: *lnpinc* is the log of per capita real permanent income, *lndinc* is per capita real disposable income, *pc* is a price deflator for total consumer expenditure, *pcncs* is the deflator

[2] An example of such a model is the 800-equation Data Resources Model (DRI) of the U.S. Economy (Eckstein O.,1983) that was widely used for economy forecasts and to evaluate various policy proposals. The equations of this model are re-estimated each year, upon availability of the national income account data, to get revised forecasts and simulation results.

for consumer expenditure on clothing and shoes, and W is the net worth of the household sector.

Although these types of models can be useful for short-term macroeconomic forecasting, Lucas (1972) presented a compelling argument, now known as the Lucas Critique, that showed they have limited value for policy analysis. In the wake of the Lucas Critique, policy evaluations are now often done using CGE models. These models are based on microeconomic foundations and thus have a strong basis in economic theory. Even though CGE models can be immune to the Lucas Critique, CGE models have their own weaknesses. The models are constructed using functional forms that severely limit the relationships between variables, and thus impose very rigid restrictions on the number of (and interaction between) independent variables in regression equations. The functional forms often impose severe restrictions on parameters in order to be consistent with theory. For example, functional forms belonging to the Constant Elasticity of Substitution (CES) class restrict substitution between all pairs of inputs to a single parameter. More flexible forms, such as the translog and generalized McFadden functions, remove this constraint but impose strong restrictions to ensure that the estimated functions have dcurvature properties consistent with optimization (Jorgensen, 1984, Diewert, and Wales, 1987).

Thus, in contrast to the econometric forecasting models, the restrictive nature of typical CGE equations results in poor econometric fit and large unexplained variance. For example, the capital share equation for energy for a

CGE model might be written as shown below.[3] The value of R^2 for this equation is only 0.559, much lower than the forecast model described above.

$$\ln(\omega_{KE}) = \ln\left(0.261\left(\frac{P_E}{P_K}\right)^{-0.149}\right)$$

This makes it particularly important for CGE results to include confidence intervals: reporting results without confidence intervals suggests that the results are known precisely, which is clearly incorrect given that the equations in the model are often a poor fit to the underlying data. Compare, for example, the following hypothetical policy statements: "Revenue from the tax will be between 660 million dollars and 740 million dollars with a 95 percent degree of confidence" versus "Revenue from the tax will be 700 million dollars." The latter statement implies an unjustified degree of precision and gives a false sense of security to the policy makers (Kendrick, 1988).

It would be surprising and unacceptable not to report standard errors for estimation in econometric studies, and so should be the case for CGE modeling, because it is an extension of econometrics. Moreover, confidence intervals would complement the traditional strength of CGE modeling–a rigorous theoretical framework–by providing an equally rigorous measure of the fit of a model's equations and the degree of precision in its parameter estimates.

---

[3] The regression equation for the capital share equation of the energy industry for the U.S. Economy described in Chapter 4. In fact the R^2 for the producer equations are much lower than the consumer equations for the model described in Chapter 4. Although it can be argued that this CGE model equation should have poor fit, it should be noted that R^2 in the consumer side equations are high because of the properties of the functional form and its relationship with income.

The remainder of the chapter presents a simple partial equilibrium model with two equations, and shows an approach that can be used to calculate confidence intervals. The equations are econometrically estimated and 95 percent confidence intervals are calculated for the dependent variables. Further, the resulting confidence intervals are compared to sensitivity analysis, which shows why sensitivity analysis is not an adequate substitute and provides motivation for the use of confidence intervals in policy analysis.

2.2 PARTIAL EQUILIBRIUM MODEL DESCRIPTION

Consider an economy in which production is represented by a single competitive firm. The firm produces an output good and sells it at its marginal cost. If the price the producer receives for the commodity is represented by p, then a linear demand equation for the good could be written as follows:

$$Q = \alpha + \beta p \tag{2.1}$$

where, Q is the output quantity and p the producer price. Parameters α and β are the intercept and the slope of the demand curve, respectively. The total revenue, R, generated by the firm for a given price and quantity is then written as:

$$R = pQ \tag{2.2}$$

Substituting the demand equation 2.1 into the total revenue equation 2.2 simplifies the model to a single equation, shown in equation 2.3:

$$R = p(\alpha + \beta p) \tag{2.3}$$

The model thus involves two variables, R and p, and two unknown parameters.

2.3 THE DATA AND ESTIMATION RESULTS

The unknown parameters α and β can be estimated from data on p and Q by adding a stochastic error term to equation 2.2:

$$Q = \alpha + \beta P + \varepsilon \tag{2.4}$$

where ε denotes a normally distributed stochastic variable with mean zero To make the discussion concrete, an example demand equation was estimated using historical data for the US service sector.[4] Estimation results and standard errors are shown in Table 2.1. The estimated parameter variance-covariance matrix, $\hat{\Sigma}$, and the estimated residual variance, s^2, are also shown below.

[4] Construction of the data set is explained in Appendix A.

Table 2.1: Parameter Estimates[5]

Parameters	Estimates	Std. Error
$\hat{\alpha}$	1.1398	0.070299
$\hat{\beta}$	-0.373	0.089472

Estimated Parameter Covariance Matrix:

$$\hat{\Sigma} = \begin{bmatrix} 0.004942 & -0.006183 \\ -0.006183 & 0.008009 \end{bmatrix}$$

Estimated Residual Variance:

$$s^2 = 0.006498$$

2.4 CONFIDENCE INTERVAL CALCULATION

In order to provide a strong foundation for the analysis later in this dissertation, it is helpful to review the general method by which confidence intervals are constructed. Suppose a variable Y depends on a predetermined variable X and an unknown parameter θ according to the relationship: $Y = f(X,\theta)$. Further, suppose that θ has been estimated by appending an additive stochastic disturbance, ε, to f and running an appropriate regression. Let the estimate be

[5] Standard error estimates appear in parentheses. Double asterisk indicates significant difference from zero at 95 percent.

denoted by $\hat{\theta}$ and its variance be denoted by σ. The variance of Y for a given X, σ_Y^2, is given by the following:

$$\sigma_Y^2 = Var\left(f(X,\hat{\theta}) + \varepsilon \right)$$

The standard error, therefore, is given by:

$$\sigma_Y = \sqrt{Var\left(f(X,\hat{\theta}) + \varepsilon \right)} \tag{2.5}$$

If f is linear in θ, Y will be normally distributed and will have a confidence interval around its mean that can be written as:

$$\hat{Y} \pm t_{crit} s_Y \tag{2.6}$$

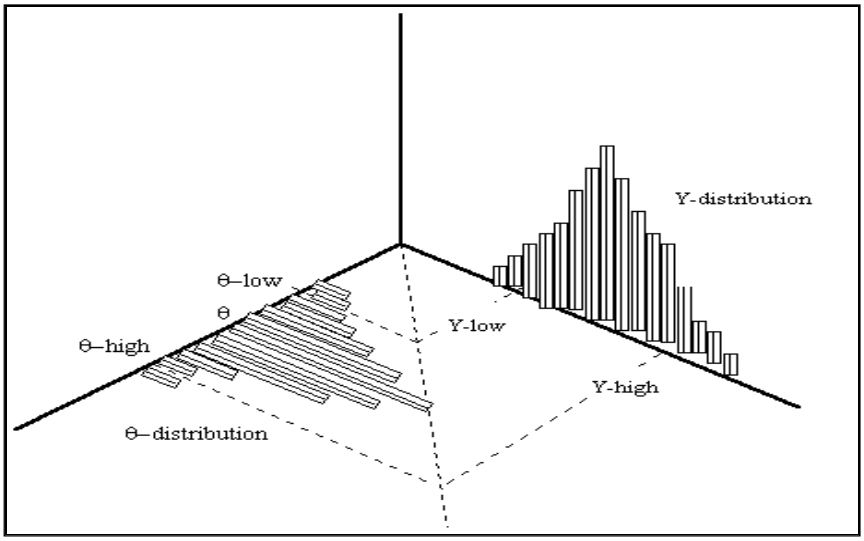
where s_Y is the estimator of σ_Y.

In Figure 2.1, the above calculation is graphically shown for a linear function that depends positively on $\hat{\theta}$. The distribution of $\hat{\theta}$ is shown by the histogram labeled "θ-distribution" and the resulting distribution of Y (for example, revenue) is shown by the histogram labeled "Y-distribution". The linear transformation function maps the parameter distribution into the Y distribution.

A more direct and simpler illustration of the variation in the total revenue is shown in Figure 2.2. The figure shows several demand curves having a fixed intercept parameter but with different slope parameters. The demand curve, *D0*,

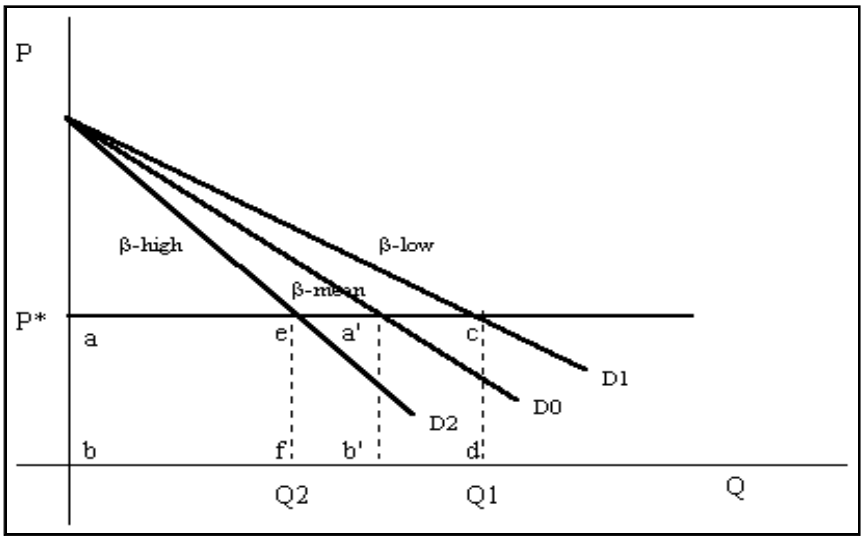
represents demand curve evaluated at the estimates of α and β parameters. Demand curves *D1* and *D2* curves correspond to demand curves for two extreme realizations of $\hat{\beta}$: β-high and β-low.

Figure 2.1: Graphical Representation of Uncertainty Transformation



The total revenue corresponding to demand curve, D0, is given by the area *aba'b'*. However, if the true value of β lies between β-low and β-high, then the total revenue could vary from area *abcd* to *abef* respectively. For example, if the true value of β is β-low, actual revenue will be unexpectedly high relative to projections based on the estimated value of β; if the true value is β-high, it will be unexpectedly low. Total revenue thus depends strongly on the value of the true slope parameter. In this simple illustration, if β-low and β-high are chosen such that they form a 95 percent confidence interval for β, then area *abef* to *abcd* represents a 95 percent confidence interval for total revenue.

Figure 2.2: Change in Total Revenue Due to Change in $\hat{\beta}$



The confidence interval for revenue given a specific, non-stochastic value of the independent variable, p_0, can be calculated as follows. Substituting equation 2.4 into equation 2.2 gives the value of revenue as a function of the price, the parameters, and the disturbance term. In particular, the predicted value of total revenue, R_f, for a known value p_0 can be written as:

$$R_f = p_0\alpha + p_0^2\beta + p_0\varepsilon_f \tag{2.7}$$

where ε_f is the random error term for the date of the prediction. Since the mean of ε_f is zero, the point prediction for total revenue for known p_0 is given by:

$$\hat{R}_f = p_0\hat{\alpha} + p_0^2\hat{\beta} \tag{2.8}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the estimates of α and β.

The difference between equations (2.7) and (2.8) gives the prediction error, e_f, which can be written as:

$$e_f = R_f - \hat{R}_f = \left(p_0\alpha + p_0^2\beta + p_0\varepsilon_f\right) - \left(p_0\hat{\alpha} + p_0^2\hat{\beta}\right) \tag{2.9}$$

Dropping the *f* subscript and taking expectations gives:

$$E(e) = -p_0 E\left(\alpha - \hat{\alpha}\right) - p_0^2 E\left(\beta - \hat{\beta}\right) + p_0 E(\varepsilon) = 0 \tag{2.10}$$

since $E(\varepsilon) = 0$ and $\hat{\alpha}$ and $\hat{\beta}$ are unbiased estimators of α and β, $\hat{R}_f$ is an unbiased predictor of revenue. The variance of the prediction error can be calculated as follows:

$$\begin{aligned}\operatorname{var}(e) &= E\left(\left(-p_0\left(\alpha - \hat{\alpha}\right) - p_0^2\left(\beta - \hat{\beta}\right) + p_0\varepsilon\right)^2\right) \\ &= p_0^2 \operatorname{var}\left(\hat{\alpha}\right) + 2p_0^3 \operatorname{cov}\left(\hat{\alpha}, \hat{\beta}\right) + p_0^4 \operatorname{var}\left(\hat{\beta}\right) + p_0^2 \operatorname{var}(\varepsilon)\end{aligned}$$

$$= \begin{bmatrix} p_0 & p_0^2 \end{bmatrix} \begin{bmatrix} \operatorname{var}\left(\hat{\alpha}\right) & \operatorname{cov}\left(\hat{\alpha}, \hat{\beta}\right) \\ \operatorname{cov}\left(\hat{\alpha}, \hat{\beta}\right) & \operatorname{var}\left(\hat{\beta}\right) \end{bmatrix} \begin{bmatrix} p_0 \\ p_0^2 \end{bmatrix} + p_0^2 \operatorname{var}(\varepsilon)$$

Writing this in matrix notation gives:

$$var(\Delta) = T' \sum T + p_0^2 \sigma^2 \tag{2.11}$$

where T is a column vector: $\begin{bmatrix} p_0 & p_0^2 \end{bmatrix}^T$ and Σ is the covariance matrix of the estimates. The covariance matrix, Σ, for the model can be expressed in terms of the variance of the disturbances as:

$$\Sigma = (X'X)^{-1} \sigma^2 \tag{2.12}$$

where matrix X is represents the sample data matrix corresponding to the explanatory variables.

Equation 2.11, cannot be used directly because σ^2 is unknown. However, σ^2 is can be replaced by its estimate, s^2:

$$s^2 = \frac{\sum e^2}{n-k}$$

which is the sum of the squared residuals divided by the degrees of freedom. The estimate of the covariance, $\hat{\Sigma}$, for this model is calculated by the regression package and is shown in Table 2.1. Substituting this into equation 2.11 gives the estimate of the variance of the prediction error:

$$T' \hat{\Sigma} T + p_0^2 s^2 \tag{2.13}$$

Since the prediction error is a sum of normal variables, it is also normally distributed. Thus, the quantity

$$\frac{R_f - \hat{R}_f}{\sqrt{T'\hat{\sum}T + p_0^2 s^2}}$$

follows a *t*-distribution with (*n*-*k*) degrees of freedom.[6] Then the overall 95 percent confidence interval for total revenue can be shown as:

$$\hat{R}_f \pm t_{0.025}\sqrt{T'\hat{\sum}T + p_0^2 s^2}$$

The expression can be expanded by eliminating $\hat{R}_f$ using equation 2.8:

$$p_0\hat{\alpha} + p_0^2\hat{\beta} \pm t_{0.025}\sqrt{T'\hat{\sum}T + p_0^2 s^2} \qquad (2.14)$$

Alternatively, it can be expressed in terms of *s*:

$$p_0\hat{\alpha} + p_0^2\hat{\beta} \pm t_{0.025}s\sqrt{T'(X'X)^{-1}T + p_0^2} \qquad (2.15)$$

[6] For arbitrarily large degrees of freedom the value of the t_{crit} is the same as that of normal distribution, which is approximately equals two at 5 percent significance level. Through out this thesis, this value of t_{crit} will be used for calculating the confidence interval.

It is interesting to note that as the sample size becomes large, the overall confidence interval asymptotically approaches:

$$p_0 \hat{\alpha} + p_0^2 \hat{\beta} \pm t_{0.025} s \sqrt{p_0^2}$$

In this situation, the confidence interval has collapsed until all that remains is the component due to the unexplained variation. This confidence interval, in other words, would apply if the parameters were estimated with complete precision. However, in practice the samples used to estimate CGE models are far too small for any given confidence interval to approach its asymptotic limit.

In many circumstances, the goal of a CGE analysis is to compare a proposed or counterfactual policy against a specified base case. In that situation, it is natural to assume that the value of the disturbance term, ε, will be invariant with respect to the policy being evaluated; that is, the disturbance will have the same realization under both policies. Put another way, the goal of a CGE study is often to determine the effect of the policy on the mean values of the model's endogenous variables. As a result, it may be more relevant to focus on the confidence interval for the mean. For example, the confidence interval for the mean value of total revenue can be derived as follows. The prediction error for the mean is:

$$e_f = E(R_f) - \hat{R}_f = (p_0\alpha + p_0^2\beta) - \left(p_0 \hat{\alpha} + p_0^2 \hat{\beta} \right)$$

Following similar algebra, a 95 percent confidence interval for the mean value of the total revenue, $E(R_f)$, can be written as:

$$p_0 \hat{\alpha} + p_0^2 \hat{\beta} \pm t_{0.025} \sqrt{T' \hat{\Sigma} T}$$

or, equivalently, as:

$$p_0 \hat{\alpha} + p_0^2 \hat{\beta} \pm t_{0.025} s \sqrt{T' (X' X)^{-1} T} \qquad (2.16)$$

The prediction variance expressed in equation 2.13 can be decomposed into two components. The first component, $T' \hat{\Sigma} T$, results from imprecision in the parameter estimates and the second component, $p_0^2 s^2$, is due to unexplained variation in the data. Thus, the confidence interval given in equation 2.14 measures the overall precision of the forecast of revenue. Equation 2.16, on the other hand, measures the precision of the forecast of *mean* revenue. To make a distinction between these two types of confidence intervals, henceforth the confidence interval given by equation 2.14 will be referred as the *overall confidence interval* and the confidence interval for the mean, given by equation 2.16, will be referred to as simply the *confidence interval*. Unless otherwise noted, all subsequent references to confidence intervals will refer to 95 percent confidence intervals.

2.5 NUMERICAL RESULTS

As an example, consider the overall confidence interval for revenue when price index p_0 is equal to its sample average value of 0.77235 (in 1982 dollars). The mean value of revenue—or the point prediction—at that price is 658 million dollars. The confidence interval for the mean is 638 to 678 million dollars; plus or minus about 3 percent of the mean. However, the overall confidence interval is much wider: 532 to 784 million dollars, or about 19 percent of the mean.

Figure 2.3: Total Revenue Confidence Intervals

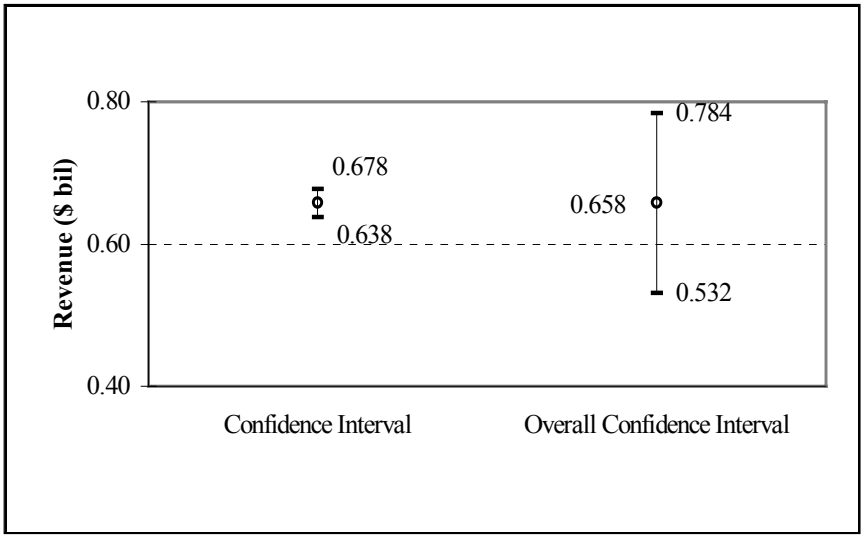


Figure 2.3 illustrates the two confidence intervals graphically. A vertical line represents each confidence interval, where the mean value lies at the center and the end points show the bounds of the confidence interval. The confidence interval for the mean is much tighter than the overall confidence interval: the overall confidence interval is about six times as wide. Although the tight

confidence interval for the mean indicates that revenue is relatively insensitive to parameter uncertainty, it alone is not an adequate measure of the precision of the model.

2.6 COMPARISON BETWEEN CONFIDENCE INTERVALS AND SENSITIVITY ANALYSIS

It is often the case in CGE models that parameters are not estimated but are calibrated to data from a benchmark year. A major drawback of this approach is that no statistical information is available for use in computing confidence intervals. In such circumstances, "sensitivity analysis" is sometimes performed to check the sensitivity of the results to the values of the parameters.

Sensitivity analysis is performed by perturbing one or more parameters from its benchmark value and then evaluating the effect of the change on the variables of interest. The range of the results is used as a measure of the robustness of the model. If the variables do not deviate too much from their benchmark values, then the model is said to be insensitive to parameter uncertainty. However, if there are large fluctuations in the results the model is said to be highly sensitive to the perturbed parameters.

In the model presented above, sensitivity analysis could be conducted by setting $\hat{\alpha}$ to some arbitrarily chosen high and low values, say 1.28 and 0.99 and then calculating revenue.[7] Revenue corresponding to the high value is \$766 million, and that corresponding to the low value is \$549 million dollars.

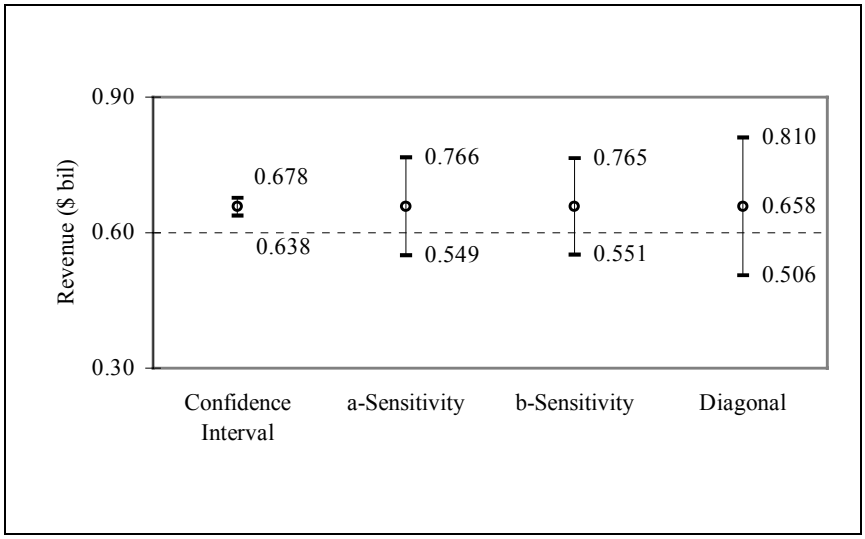
[7] In order to compare the results between sensitivity analysis and the confidence interval, high and low values of α–parameter are chosen at ± 2 standard deviations from its estimated value.

Similarly, if $\hat{\beta}$ is perturbed to -0.55 and -0.19, revenue changes from the mean value to $551 million and $765 million, respectively.

Figure 2.4 shows a comparison of the confidence intervals and the sensitivity analysis results. The confidence interval for mean revenue is labeled "Parameter" (because it results from uncertainty in the parameters estimates). The standard sensitivity analysis ranges determined by perturbing $\hat{\alpha}$ and $\hat{\beta}$ individually are shown as "a-Sensitivity" and "b-Sensitivity". In addition, the figure also includes an interval labeled “Diagonal”. It is a confidence interval for the mean calculated using a modified variance-covariance matrix for the parameter in which the off-diagonal elements were set to zero. It shows the importance of cross correlations in the parameter estimates: ignoring the interdependence between the estimates, as is done in the “Diagonal” case, causes the confidence interval for mean revenue to become substantially larger. This point will be discussed in more detail below.

Figure 2.4 shows that the range of total revenue under sensitivity analysis results for a-Sensitivity and b-Sensitivity exceeds the length of the confidence interval by approximately a factor of five. The value of revenue could be as large as 766 million dollars or small as 559 million dollars. As will be explained in detail below, the fact that sensitivity analysis produces a wider interval (and hence a more pessimistic view of the model’s precision) is an artifact of the particular example used here. The general lesson from Figure 2.4 is that the results of sensitivity analysis can be a very poor approximation to a variable’s true confidence interval.

Figure 2.4: Comparison of Confidence Intervals and Sensitivity Analysis Range



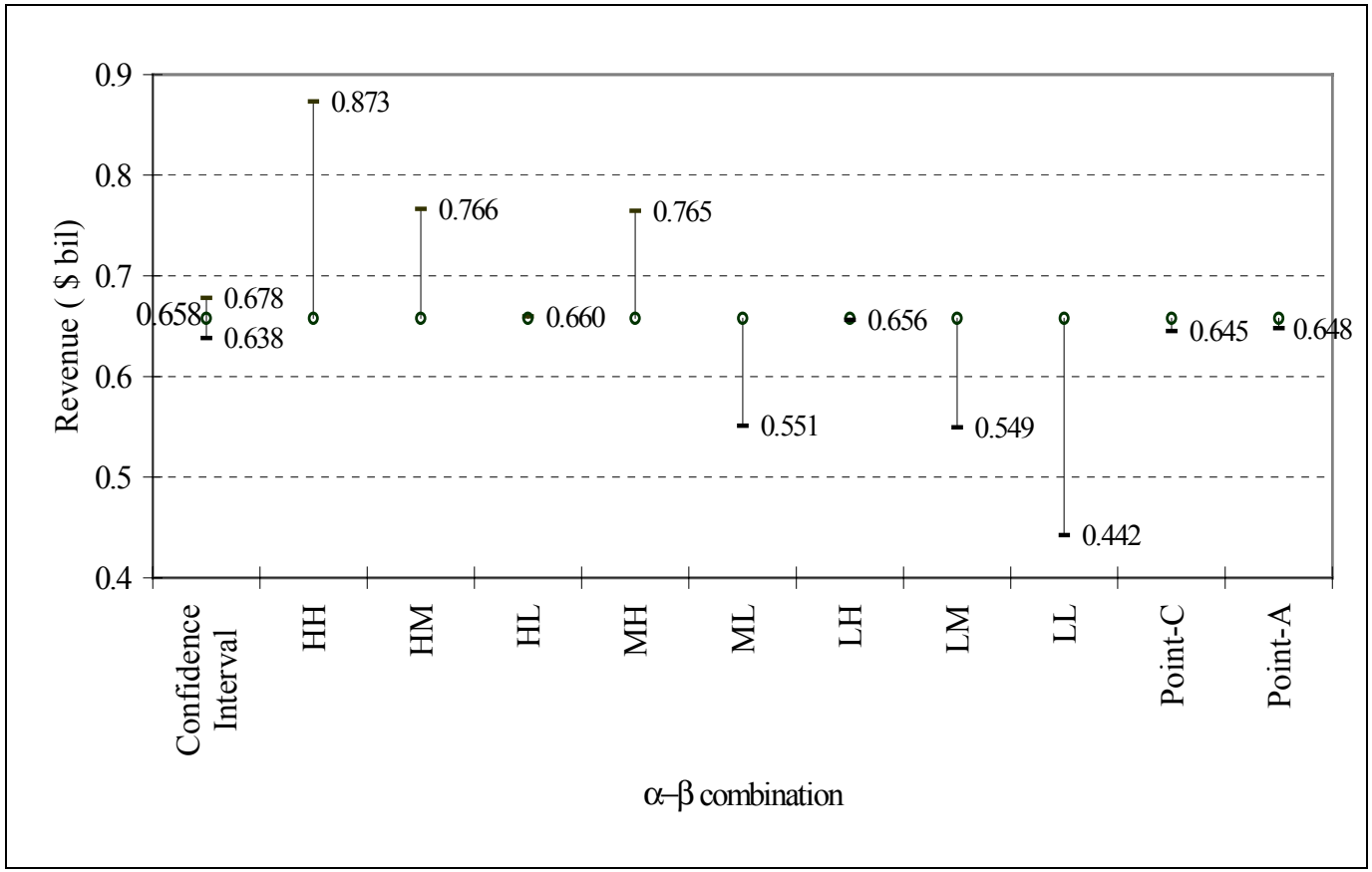
To explore the relationship between sensitivity analysis and confidence intervals in more detail, consider carrying out a sensitivity analysis by making simultaneous perturbations in both parameters. Suppose parameter is set to three different values: a high value (H), its mean value (M), and a low value (L). The high and low values are chosen to two standard errors from the mean. In all, eight pairs of the parameter combinations can be evaluated (there are nine potential combinations but one uses the means for both parameters and is thus the model itself). The pairs could be named HH, HM, HL, MH, ML, LH, LM, and LL where the first letter of the pair represents the value taken by $\hat{\alpha}$ and the second represents the $\hat{\beta}$ value. For example, category HH represents high values of both $\hat{\alpha}$ and $\hat{\beta}$ while category HM represents the high value of $\hat{\alpha}$ and the actual estimate of $\hat{\beta}$. The Table 2.2 lists the values for all eight pairs.

Table 2.2: Parameter Values for Sensitivity Analysis

Category	$\hat{\alpha}$	$\hat{\beta}$
HL	1.280	-0.552
HM	1.280	-0.373
HH	1.280	-0.194
ML	1.139	-0.552
MH	1.139	-0.194
LL	0.999	-0.552
LM	0.999	-0.373
LH	0.999	-0.194

Figure 2.5 shows sensitivity analysis results for all the parameter combinations. Under the HL and LH parameter combinations, the range of the total revenue is very small. In fact, it falls within the true confidence interval shown under "Parameter". A sensitivity analysis restricted to these two cases would conclude that revenue is precisely determined and robust to changes in parameter values. However for the other six pairs (HH, HM, MH, ML, LM, and LL), the range produced by sensitivity analysis is much larger than the previous two cases. An analysis restricted to these six cases would conclude that revenue is strongly affected by changes in parameters—the exact opposite of the conclusion from the HL and LH cases. Which conclusion is correct depends on the relative likelihood of the alternative scenarios which is entirely outside the scope of sensitivity analysis. Thus, an important weakness of sensitivity analysis is that its conclusions depend heavily on the particular choice of perturbations made to the parameters, but the methodology provides no systematic guidance on how perturbations should be selected.

Figure 2.5: Comparison of Sensitivity Analysis Results for Different $\hat{\alpha}$ and $\hat{\beta}$ Combinations



Because sensitivity analysis does not take into account the likelihood of alternative perturbations it cannot distinguish between combinations of parameters that are plausible and those that not plausible. For example, a researcher carrying out the sensitivity analysis above has no way of knowing the likelihood of the HH case (α and β are 1.28 and -0.19) being true (or of any other $\hat{\alpha}$ and $\hat{\beta}$ combination, for that matter). Without statistical information on the joint distribution of the parameter estimates, it is not possible to determine—for example—whether HH is more or less likely to be true than HL.

In this example, however, statistical information *is* available on the distribution of the estimates. The joint distribution for the model parameters can be used to compute the likelihood of different parameter combinations.[8] A convenient way to summarize the distribution is by constructing a 95 percent confidence ellipsoid for the true values of the parameters. The confidence interval can be found from the following expression:

$$\frac{1}{2}\left(\alpha-\hat{\alpha},\beta-\hat{\beta}\right)^{T}\left(\hat{\Sigma}\right)^{-1}\left(\alpha-\hat{\alpha},\beta-\hat{\beta}\right)=F(2,37)_{0.95} \qquad (2.17)$$

In equation 2.17 everything is known except α and β. The values of α and β that satisfy the equation form a two-dimensional ellipsoid as shown in Figure 2.6.

An interesting feature of the parameter distribution is that the confidence ellipsoid is very narrow and lies along a downward-sloping line (the latter

[8] A joint F-distribution is assumed for the parameters.

reflecting the negative correlation between the parameter estimates). From the figure, it is clear that there are many combinations of parameters—some quite close to the estimated values—that are extremely unlikely to be true given the observed data. For example, the HH combination used in the sensitivity analysis (α and β equal to 1.28 and -0.19) is very, very far outside the ellipsoid. However, the LH combination (α and β equal to 1 and -0.19) could be inside the contour.

The location of all eight of the sensitivity analysis combinations relative to the ellipsoid is shown in Figure 2.7. In addition, iso-revenue loci were generated for various levels of revenue; these are shown in the figure by dotted lines. The result is a graph that illustrates the likelihood associated with each revenue point. The iso-revenue locus with revenue equal to the mean is indicated by line-66 in the figure.

Sensitivity analysis cases HL and LH lie inside the confidence ellipse, indicating that such parameter combinations cannot be rejected. These points also lie very close to the mean iso-revenue line (line-66), resulting in a tight sensitivity analysis range. The parameter combinations corresponding to the other extreme cases, HH and LL, are outside the ellipse and could be rejected with a high degree of confidence. It is interesting to note that the HH and LL points lie far from the mean iso-revenue line. Under these realizations (HH and LL) the revenue interval generated by sensitivity analysis is deceptive: it is large but has essentially zero probability of occurrence.

Figure 2.6: 95 Percent Confidence Interval Parameter Ellipsoid

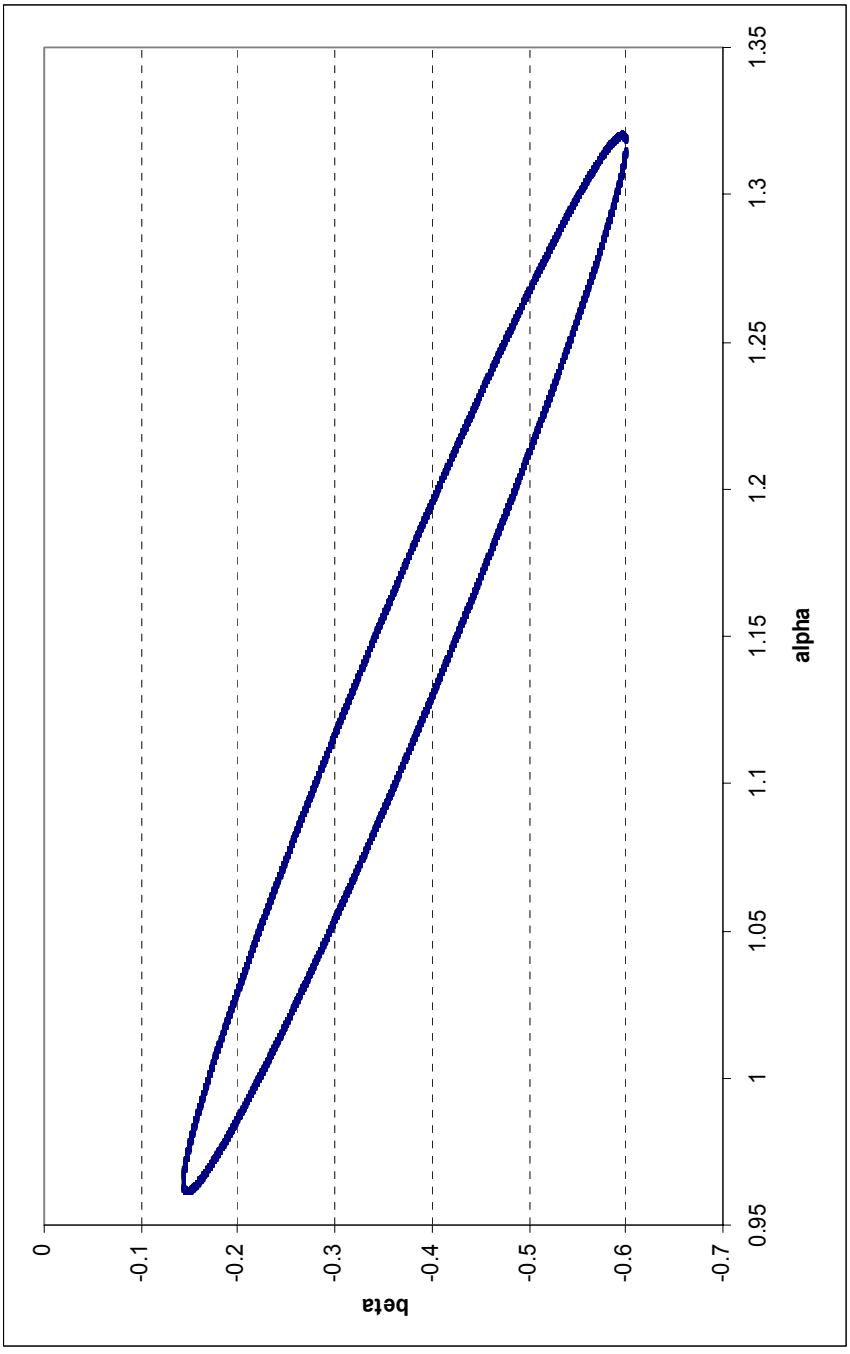
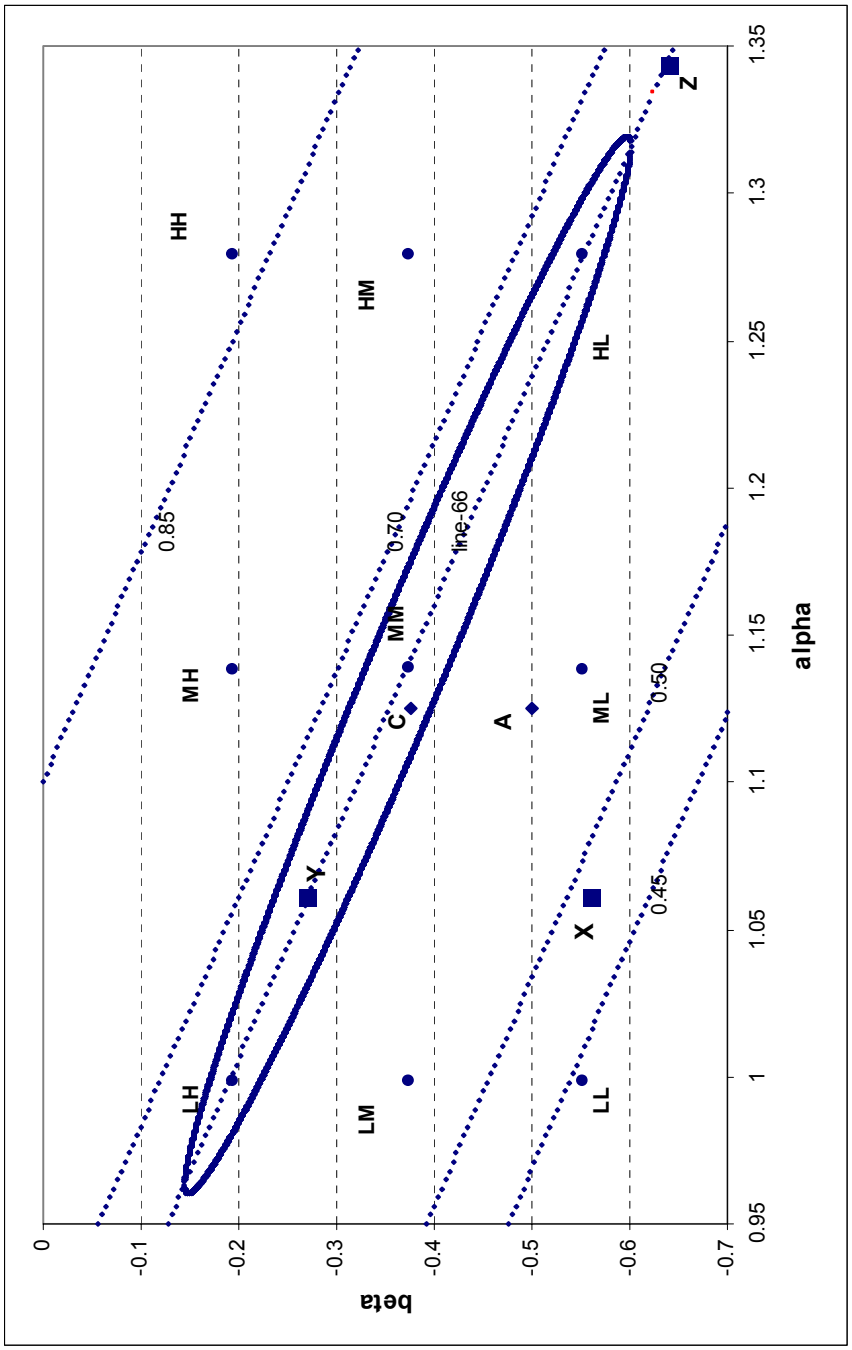


Figure 2.7: 95 Percent Confidence Interval Parameter Ellipsoid and Iso-revenue Lines



It is possible that parameter combinations that are close to the mean iso-revenue line but outside the ellipsoid would result in a tight sensitivity analysis interval but despite having a low probability of occurrence. For instance, point Z is far from the estimated coefficients (point MM) but just happens to be close to the mean iso-revenue line. A sensitivity analysis comparing the effects of MM and Z on revenue would, therefore, dramatically understate the model's sensitivity. It would show essentially no change in revenue despite a considerable change in the two parameters. The model would thus appear to be insensitive to its parameter values when in fact the point of comparison, Z, just happened to have the same revenue as MM. This exercise would give the false impression that revenue is insensitive to changes in the parameter values.

From Figure 2.7 it is clear that sensitivity analysis using parameter vectors along the mean iso-revenue line would result in a tight range (little sensitivity), whereas one using vectors away from the iso-revenue line would show a considerably larger range. The weakness of sensitivity analysis is that it provides no way to choose appropriate combinations of perturbations to the parameters. In essence, it involves permuting the values of the parameters without any knowledge about their joint distribution. An alternative way to see why sensitivity analysis can give misleading conclusions is presented below.

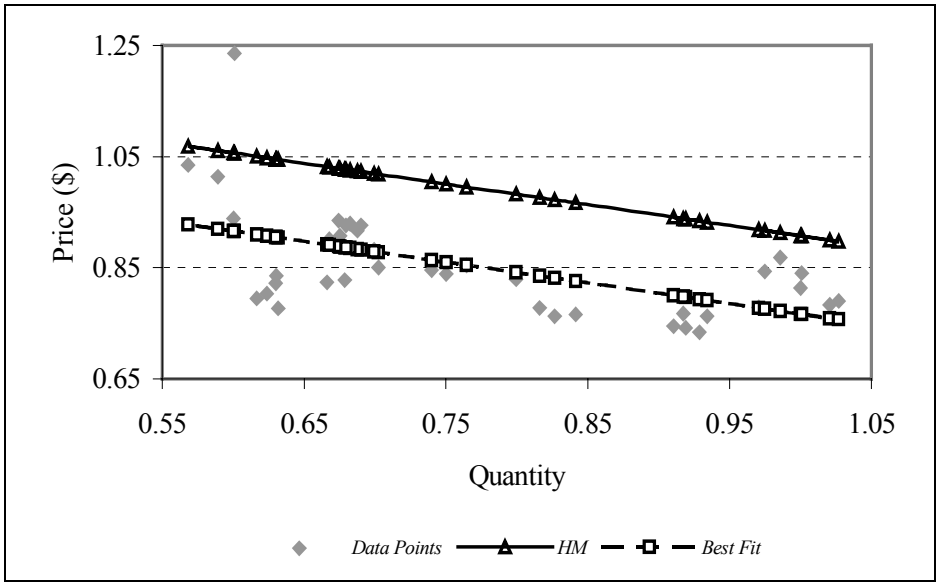
2.7 SENSITIVITY ANALYSIS AND DATA

Another way to look at sensitivity analysis is to compare the functions it generates with the actual data, since data is available for this example. Figure 2.8

shows a scatter plot of the service sector data. The least squares regression line is also shown, along with a line corresponding to the HM sensitivity analysis case. The HM line is shifted upward relative to the regression and does not fit the data. Performing this sort of sensitivity analysis, in other words, is equivalent to imposing a relationship that is inconsistent with the data.

A similar pattern holds true when only the $\hat{\beta}$ parameter is perturbed. In Figure 2.9, the ML sensitivity analysis case is compared with the data and the least squares regression line. Again, the ML case clearly does not fit the data.

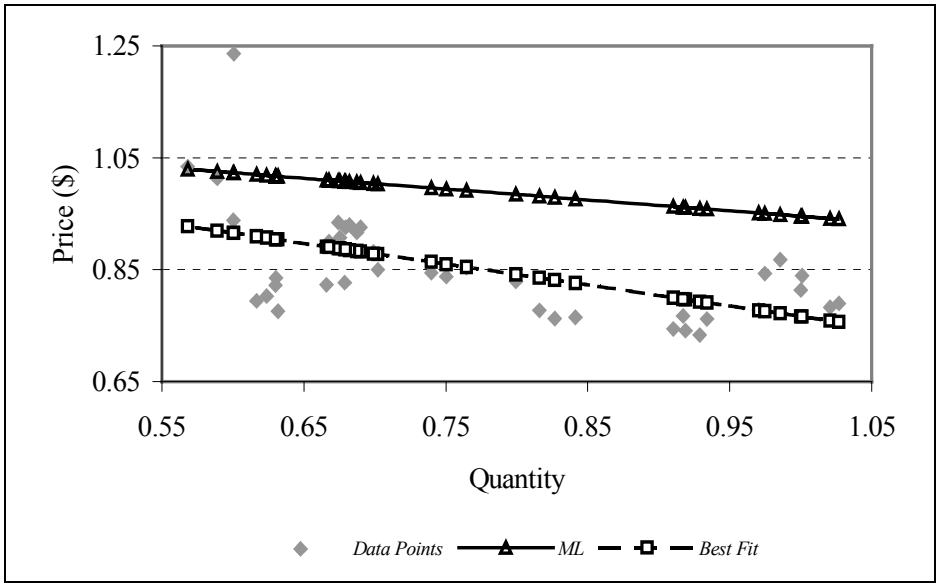
Figure 2.8: Sensitivity Analysis Lines for Change in $\hat{\alpha}$



The gap between sensitivity analysis and the underlying data can be even larger if both parameters, $\hat{\alpha}$ and $\hat{\beta}$, are perturbed at the same time and in the same direction. Line-HH, for example, is shown in Figure 2.10. It lies outside

almost every one of the data points. On the other hand, if the parameters are perturbed in opposite directions (cases HL and LH, for example), then the demand lines corresponding to it is very close to the regression line.

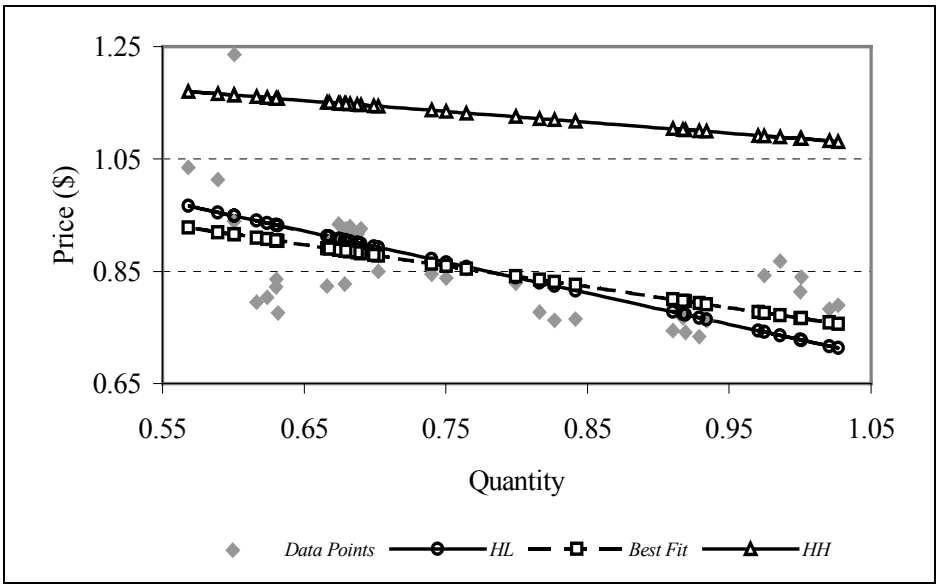
Figure 2.9: Sensitivity Analysis Lines for Change in $\hat{\beta}$



Together, Figures 2.8, 2.9, and 2.10 illustrate the fact that—in this model—when only one parameter is perturbed, it is almost always the case the sensitivity analysis generates an incorrect range. However, the range is tighter if the parameters are perturbed in the opposite direction: since $\hat{\alpha}$ and $\hat{\beta}$ are negatively correlated, increasing or decreasing one parameter demands that the

other be decreased or increased, respectively, if the perturbed equation is to remain close to the original data.[9]

Figure 2.10: Sensitivity Analysis Lines for Change in $\hat{\alpha}$ and $\hat{\beta}$



In sum, the ad hoc process of perturbing of the parameters in standard sensitivity analysis fails to take into account the correlation between estimated parameters. It can be similar to constructing a confidence interval using only the diagonal elements of the parameter variance-covariance matrix (see the "Diagonal" entry in Figure 2.4). This arises because sensitivity analysis is unable to account for cross correlation between the parameter estimates. Thus, it is

[9] The correlation coefficient between α and β parameter is -0.983. The negative value indicates that the perturbation has to be in opposite direction for the analysis to be consistent. However, for models with more parameters, the direction of perturbation is difficult to determine.

imperative to use the full variance-covariance matrix and calculate a confidence intervals whenever possible.

2.8 CHAPTER SUMMARY

The values of the variables in any model are never known with perfect certainty. Thus, it is important to quantify and report the precision of the parameter estimates used in a model, and the precision of the model's results. This chapter presented and compared two basic approaches for measuring the uncertainty in a model's output: confidence intervals and sensitivity analysis. As an example, a model based on a simple revenue equation with two parameters was econometrically estimated for the US service sector. The overall confidence interval (including the unexplained residual) and the confidence interval for the mean were calculated. The confidence interval, which accounts for parameter uncertainty, was found to be much tighter than the overall confidence interval.

Further, the confidence intervals were also compared to the ranges of variation determined by sensitivity analysis. Different guesses of the parameter combinations produced distinctly different ranges. These results could lead one to infer misleading conclusions about the sensitivity of the model to its parameters.

Finally, the analysis of confidence intervals was extended graphically and it was shown that the failure of sensitivity analysis to incorporate realistic cross correlations between parameters is a key reason why confidence intervals are superior to sensitivity analysis.

Chapter 3: Confidence Intervals for a Small Static CGE Model

3.1 INTRODUCTION AND MODEL ECONOMY

In this chapter a small static computable general equilibrium model, henceforth referred as the PROTO model, is designed, and the procedure for calculating confidence intervals is elaborated. The PROTO model consists of a single representative consumer, a producer, and a government sector. The equations for each of the agents are outlined in the sections below. Since PROTO is only an illustrative model, it does not attempt to represent to a real economy. However, the model contains all the relevant structures. The parameters of the model are econometrically estimated for the US economy.

The model is simple enough to illustrate each step of the method in a tractable algebraic form. This chapter presents the linearization process and the confidence interval calculations in explicit detail. The chapter gives a working methodology for evaluating confidence intervals that can be performed routinely without significant computational burden. The procedure described in this chapter will be the basis for calculating intertemporal confidence intervals in the chapters to follow.

3.1.1 Producer Behavior

A single producer that produces a single composite output, Q, represents the production side of the economy. The output is consumed by the household and also used as an intermediate input in production. The producer demands

intermediate goods, X, and two primary factors of production: capital, K_d, and labor, L_d, as inputs. The producer price of output is P and the purchaser's prices of capital services and labor are represented by P_k and W, respectively. The producer seeks to minimize total cost subject to a technology constraint represented by a Cobb-Douglas production function:

$$Q = AX^{\alpha} K_d^{\beta} L_d^{1-\alpha-\beta} \tag{3.18}$$

where A, α, and β are production function parameters. Total cost can be written as:

$$PX + P_K K_d + WL_d$$

The producer's problem then can be written as:

$$\begin{aligned} &min \quad PX + P_K K_d + WL_d \\ &subject\ to \\ &\qquad Q = AX^{\alpha} K_d^{\beta} L_d^{1-\alpha-\beta} \end{aligned} \tag{3.19}$$

The Lagrangian for the above problem can be formulated as:

$$L = PX + P_k K_d + WL_d + \lambda\left\{Q - AX^{\alpha} K_d^{\beta} L_d^{1-\alpha-\beta}\right\} \tag{3.20}$$

The input demand equations below can be derived from the first order conditions:

$$PX = \alpha PQ \tag{3.21}$$

$$P_K K_d = \beta PQ \tag{3.22}$$

$$\mathrm{WL_d} = (1 - \alpha - \beta)\mathrm{PQ} \tag{3.23}$$

3.1.2 Consumer Behavior

The representative household consumes an aggregate consumption good, *C*. The household receives income by providing labor and capital services to the producer and in addition receives a government transfer in the form of a lump-sum subsidy, S. The household also decides on the amount of labor to supply (and thus leisure to consume). The household owns the total capital stock, $\overline{K}$, which is exogenous. The total time endowment is fixed at *H*.

Consumer behavior is represented by a Cobb-Douglas utility function of consumption and leisure, *J*:

$$U = C^{\gamma} J^{1-\gamma} \tag{3.24}$$

where γ is utility parameter. The household maximizes utility subject to its budget and time constraints.[10] The consumer problem can be written as:

[10] As said at the outset that this is only an illustrative example. Cobb-Douglas specification is used for simplifying the model specification.

$$
\begin{aligned}
&max \quad C^{\gamma} J^{1-\gamma} \\
&subject\ to \\
&\qquad \hat{P} C = W L_s + P_k K_s + S \\
&\qquad H = L_s + J \\
&\qquad \hat{P} = P(1+t)
\end{aligned}
\tag{3.25}
$$

where $\hat{P}$ is the consumer price inclusive of the tax, *t*.

The first order conditions for the household's problem can be used to derive the behavioral equations below:

$$PC(1+t) = \gamma Y \tag{3.26}$$

$$WJ = (1 - \gamma)Y \tag{3.27}$$

$$Y = WH + P_k K_s + S \tag{3.28}$$

$$H = L_s + J \tag{3.29}$$

3.1.3 The Rest of the Model

The government sector takes a passive role in this model. It runs a balanced budget by giving back the total tax collected to the consumer in the form of the subsidy. The government budget equation can then be represented as:

$$tPC = S \tag{3.30}$$

At equilibrium, capital and labor demand and supply must balance. In addition, total output of the good must equal the sum of the demands for intermediate goods and consumption goods. Thus, the market clearing equations for the goods and factors can be written as:

$$K_d = K_s \tag{3.31}$$

$$L_d = L_s \tag{3.32}$$

$$Q = X + C \tag{3.33}$$

3.1.4 Solving the Model

The full model is represented by the thirteen equations below, which have thirteen endogenous variables: Q, X, C, J, Y, S, K_d, K_s, L_d, L_s, P, P_k, W, two exogenous variables: H, t, and four parameters: α, β, γ, and A:

$$PX = \alpha PQ \tag{3.34}$$

$$\mathrm{P_K K_d} = \beta \mathrm{PQ} \tag{3.35}$$

$$\mathrm{WL_d} = (1 - \alpha - \beta)\mathrm{PQ} \tag{3.36}$$

$$Q = AX^{\alpha} K_d^{\beta} L_d^{1-\alpha-\beta} \tag{3.37}$$

$$PC(1+t) = \gamma Y \tag{3.38}$$

$$WJ = (1 - \gamma)Y \tag{3.39}$$

$$Y = WH + P_k K_s + S \tag{3.40}$$

$$H = L_s + J \tag{3.41}$$

$$K_s = \overline{K} \tag{3.42}$$

$$tPC = S \tag{3.43}$$

$$K_d = K_s \tag{3.44}$$

$$L_d = L_s \tag{3.45}$$

$$Q = X + C \tag{3.46}$$

Eliminating equations and variables can further reduce the equations of the model. The capital and labor demand and supply variables can each be represented by a single variable, allowing two equations to be dropped. For example, capital demand, K_d, and capital supply, K_s, can be replaces by capital, K. This reduces the model to eleven equations with eleven endogenous variables.

Furthermore, the goods market clearing equation can also be dropped because of Walras Law, and the output price, P, is used as the model's numeraire and normalized to one. The model is closed and simplifies to ten equations with ten unknowns, and has three exogenous variables and four parameters. The complete final model is shown below.

$$\mathrm{PX} = \alpha \mathrm{PQ} \tag{3.47}$$

$$\mathrm{P_K K} = \beta \mathrm{PQ} \tag{3.48}$$

$$\mathrm{WL} = (1 - \alpha - \beta)\mathrm{PQ} \tag{3.49}$$

$$\mathrm{Q} = \mathrm{AX}^{\alpha} \mathrm{K_d}^{\beta} \mathrm{L}^{\,1-\alpha-\beta} \tag{3.50}$$

$$PC(1 + t) = \gamma Y \tag{3.51}$$

$$WJ = (1 - \gamma)Y \tag{3.52}$$

$$\mathrm{Y} = \mathrm{WH} + \mathrm{P_k K} + \mathrm{S} \tag{3.53}$$

$$H = L + J \tag{3.54}$$

$$K = \overline{K} \tag{3.55}$$

$$tPC = S \tag{3.56}$$

3.2 DATA AND ESTIMATION

3.2.1 Data Source

The data used for the model was obtained from a time-series of US input-output tables. The details of the data and the aggregation procedure are described in Appendix A.

3.2.2 Estimation of Parameters

The model has one consumer, γ, and three producer parameters, α, β and A. The equations used for estimation are derived by adding a stochastic term to three of the equations and appending a multiplicative term the production function to produce the four equations below:

$$PX = \alpha PQ + \varepsilon_1 \tag{3.57}$$

$$P_K K_d = \beta PQ + \varepsilon_2 \tag{3.58}$$

$$PC(1+t) = \gamma Y + \varepsilon_3 \tag{3.59}$$

$$Q = AX^{\alpha} K_d^{\beta} L_d^{1-\alpha-\beta} e^{\varepsilon_4} \tag{3.60}$$

where ε_1, ε_2, ε_3, and ε_4 are the disturbance terms.

If the error terms are assumed to be independent of each other then the parameters α, β, and γ can be estimated separately by applying ordinary least squares to the first three of these equations. The resulting estimates of α, β, and γ are 0.492, 0.18, and 0.243 respectively, as shown in Table 3.1. Although the standard error for each of the estimates can be determined by this procedure, separate estimation imposes a diagonal variance-covariance matrix for both the parameter estimates and the error terms. As pointed out in the previous chapter, information about cross correlations between parameter estimates can be important for determining confidence intervals.

In order to relax the restrictions on the variance-covariance matrices, the equations were estimated as a system using full-information maximum likelihood (FIML). FIML estimates of the parameters are shown in Table 3.1.

Table 3.1: Parameter Estimates

Parameter	OLS Estimate	FIML Estimate
$\hat{\alpha}$	0.491709	0.48876
	(0.198794E-02)	(0.32133E-02)
$\hat{\beta}$	0.180020	0.17814
	(0.15690E-02)	(0.35104E-02)
$\hat{\gamma}$	0.243751	0.21600
	(0.867605E-02)	(0.77393E-02)
$\hat{A}$	-	2.81234
		(0.01646E-02)

*Standard error are indicated within the parenthesis

The variance-covariance matrices for the estimates and the residuals are given in Tables 3.2 and 3.4, and the corresponding correlation matrices are given in Tables 3.3 and 3.5 respectively. The correlation matrix for the estimates indicates that $\hat{\alpha}$ is highly correlated with the other three estimated parameters. It is negatively correlated with $\hat{\beta}$ and $\hat{A}$ and positively correlated with $\hat{\gamma}$. Similarly $\hat{\beta}$ is negatively correlated with $\hat{\gamma}$ and positively correlated with $\hat{A}$. Likewise, $\hat{\gamma}$ is negatively correlated with $\hat{A}$. In short, all the estimates of the model are highly correlated with each other.

The residual correlation matrix in Table 3.5 indicates that the error term of the capital demand equation is positively correlated with the consumer behavior equation and negatively correlated with the technology equation. The error term in the consumer behavior equation is also negatively correlated with the error term in the technology equation.

Strong correlation amongst the parameter estimates and residuals suggests that the cross correlation effects are relevant and significant. This information would be absent from the analysis if the parameters were estimated assuming independence of the disturbances, or if the parameter variance-covariance matrix was assumed to be diagonal. In the contour diagram for the correlation matrices, Figures 3.1 and 3.2, large number of off diagonal dark spots would indicate significant cross correlation.

Table 3.2: PROTO Estimated Parameter Covariance Matrix

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{A}$
$\hat{\alpha}$	1.0325E-05			
$\hat{\beta}$	-1.0932E-05	1.2323E-05		
$\hat{\gamma}$	2.2071E-05	-2.3924E-05	5.9897E-05	
$\hat{A}$	-5.1634E-05	5.7596E-05	-1.1609E-04	2.7109E-04

Table 3.3: PROTO Estimated Parameter Correlation Matrix

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{A}$
$\hat{\alpha}$	1.0000			
$\hat{\beta}$	-0.9691	1.0000		
$\hat{\gamma}$	0.8875	-0.8806	1.0000	
$\hat{A}$	-0.9759	0.9965	-0.9111	1.0000

Table 3.4: PROTO Estimated Residual Covariance Matrix

	Eq-1	Eq-2	Eq-3	Eq-4
Eq-1	1.1544E+09			
Eq-2	-6.590E+08	7.1050E+08		
Eq-3	2.3804E+09	2.3803E+09	1.0958E+11	
Eq-4	-45.57	-3.57	-1383.38	2.095E-05

Table 3.5: PROTO Estimated Residual Correlation Matrix

	Eq-1	Eq-2	Eq-3	Eq-4
Eq-1	1.00000			
Eq-2	-0.3825	1.0000		
Eq-3	0.0363	0.9091	1.0000	
Eq-4	-0.0637	-0.8973	-0.9996	1.0000`

Figure 3.1: Contour Diagram of PROTO Estimated Parameter Correlation Matrix for FIML Estimation

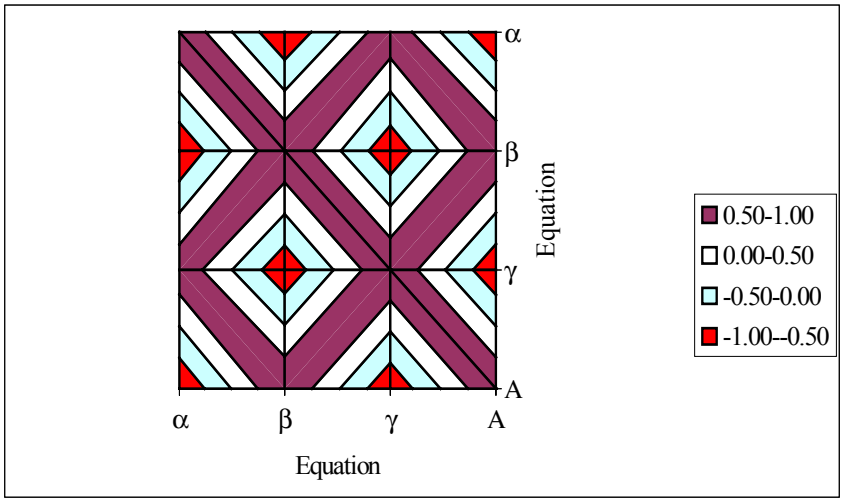
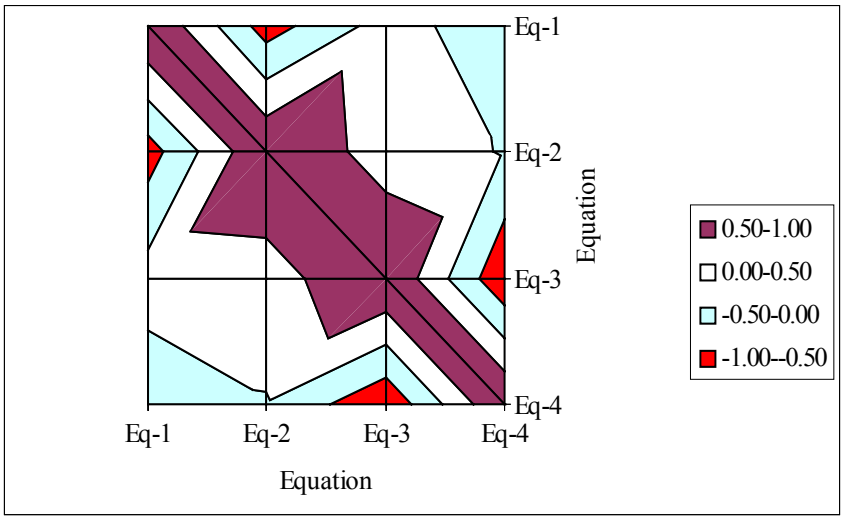


Figure 3.2: Contour Diagram of PROTO Estimated Residual Correlation Matrix for FIML Estimation



3.2.3 Exogenous Variables

The model includes three exogenous variables: the total endowment of time, H, the capital stock, $\overline{K}$, and the tax rate, t. The values of these variables were set to their 1982 levels in the dataset.[11] The aggregation process for constructing these variables is given in Appendix B. The key policy variable, the tax on the consumption good, was fixed at 10 percent for the base case simulation.[12]

3.3 METHOD FOR COMPUTING CONFIDENCE INTERVALS

In this section, a method for computing confidence intervals for the model is described. Each step of the procedure is illustrated using equations from the PROTO model.

Previous chapters noted that the overall uncertainty in a CGE model's endogenous variables arises from two sources: imprecision in the parameter estimates and the residuals in the regression equations. Using this information to compute a CGE model's confidence intervals is, in general, nontrivial. Many of the model's equations will be nonlinear and there is no general analytic method for computing confidence intervals for nonlinear models. For very small models, confidence intervals could be calculated using Monte-Carlo simulation (discussed in more detail below). However, that approach would quickly become unwieldy

[11] The PROTO model sets the year 1982 under the assumption that the model is at a steady state in 1982.
[12] This is an arbitrary chosen tax rate.

as model size increases and would be completely impractical for large models were obtaining even a single solution can be a significant computational burden.

This chapter presents an alternative procedure that uses linearization to keep the confidence interval calculation tractable, even for very large models.[13] The following section will describe the linearization process. After that, the general methodology to compute the overall confidence interval will be explained and followed by an application of the methodology to the PROTO model. Following that, it will be shown how the methodology can be used to calculate confidence intervals for the mean values of the endogenous variables. This approach will then be applied to PROTO as well.

3.3.1 Johansen Linearized Model

The first step is to linearize the nonlinear system of equations. The nonlinear system of equations for the model can be written in matrix form as:

$$F(m, n, \theta, \varepsilon) = 0 \tag{3.61}$$

where

m is p x 1 vector of endogenous variables,

n is q x 1 vector of exogenous variables,

θ is r x 1 vector of parameters,

ε is a vector of stochastic disturbances.

[13] The approach described here is fundamentally similar to the "delta method" used by statistical packages to compute standard errors for nonlinear models.

By the implicit function theorem, the solution will be a vector of endogenous variables whose values are a given by a vector function *g*:

$$m = g(m, n, \theta, \varepsilon)$$

Differentiating this expression and dividing the difference terms by the values of the corresponding variables produces an approximation to the model expressed in terms of percentage changes in the variables and parameters. The general expression can be written in matrix form as:

$$M\tilde{m} + N\tilde{n} = O\tilde{\theta} + D\Delta_{\varepsilon} \tag{3.62}$$

where $\tilde{m}$ and $\tilde{n}$ are vectors of percentage changes in endogenous and exogenous variables and $\tilde{\theta}$ is a vector of percentage changes in the parameters. The disturbance term is expressed in deviation form as Δ_{ε} and is not converted to a percentage change since each element has mean zero. Matrices *M*, *N* and *O* are first-order approximations to the model's Jacobian matrix (evaluated at the base case point around which the equations were linearized); each has *p* rows, one for each of the model's equations. *M* has *p* columns corresponding to the model's endogenous variables; *N* has *q* columns corresponding to the model's exogenous variables; *O* has *r* columns corresponding to the model's parameters; and *D* has *t* columns, where *t* is the number of equations having residuals in the estimation. *D*

is a matrix of ones and zeros mapping the residuals to the estimating equations: element D_{ij} is one if equation *i* includes residual *j* and is zero otherwise.

The vector of percentage deviations for the endogenous variables, $\tilde{m}$, is expressed relative to the vector of base case values, $\overline{m}$, and is defined as follows:

$$\tilde{m} = \frac{m - \overline{m}}{\overline{m}} \tag{3.63}$$

A similar expression is used to define $\tilde{n}$. $\tilde{\theta}$ is a vector of percent deviations in the parameters from their estimated values, $\overline{\theta}$:

$$\tilde{\theta} = \frac{\theta - \overline{\theta}}{\overline{\theta}} \tag{3.64}$$

The vector of residual variables, Δ_ε, is expressed in deviation form rather than in percentage change since the mean value of each residual is zero.[14]

---

[14] This approach is known as Johansen's method of linearization and is based on three simple rules: the product rule, the power rule, and the sum rule. These rules are illustrated below. Let $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ represent percent changes in the corresponding variables A, B, and C, and let α be a parameter.

Product rule: $A = \alpha BC$

$$\tilde{A} = \tilde{B} + \tilde{C}$$

Power rule: $A = B^{\alpha}$

$$\tilde{A} = \alpha \tilde{B}$$

Sum rule: $A = B + C$

$$\tilde{A} = \frac{B}{A}\tilde{B} + \frac{C}{A}\tilde{C}$$

One advantage of using Johansen's approach is that the partition of the model's variables into endogenous and exogenous sets can easily be modified to accommodate the needs of the particular policy analysis being conducted. This flexibility enables the approach to perform multiple policy simulations without much difficulty.[15] The disadvantage of this method is that it introduces truncation error because it is based on a first-order Taylor Series approximation.[16] Truncation error will be minimal for small changes in the model's variables. For larger changes, truncation error can be controlled to any desired degree by breaking the change into a number of smaller changes and then simulating the changes one at a time, updating the model's database between each one. The effect of the large change will be equal to the cumulative effects of the small changes. The GEMPACK software package is especially suited to performing such simulations.

For a clear exposition of how Johansen's linearization method works, consider the following equation from PROTO:

$$P_K K = \beta PQ \tag{3.31}$$

Totally differentiating the above expression gives the following equation in difference form:

---

for more detail on Johansen Linearation see chapter 3 in Dixon, Parmenter, Powell, and Wilcoxen (1989).

[15] Policy simulation is not performed with the PROTO model. GEMPACK routine SAGEM is used to perform multiple simulations, which is elaborated in chapter- 5.

[16] The higher order terms of Taylor Series expansion represents the truncation error and it depends upon the order of the step size. Simulation exercises in this thesis are conducted minimize the severity of this problem.

$$P_K dK + K dP_K = \beta P dQ + \beta Q dP + PQ d\beta \tag{3.65}$$

where dK, dP_k, dQ, and dP are variables in difference form. Multiplying and dividing each element on both sides of the equation by the variable corresponding to each differential produces the percentage change form:

$$P_K K \frac{dK}{K} + P_K K \frac{dP_K}{P_K} = \beta QP \frac{dQ}{Q} + \beta PQ \frac{dP}{P} + PQ\beta \frac{d\beta}{\beta} \tag{3.66}$$

The original equation can be used to eliminate $P_K K$ and βQP to produce the equation below (shown in derivative form and also in percentage change variables indicated tildes). Note that this equation can also be deduced by straightforward application of the Johansen's product rule.[17]

$$\frac{dK}{K} + \frac{dP_K}{P_K} = \frac{dQ}{Q} + \frac{dP}{P} + \frac{d\beta}{\beta} \tag{3.67}$$

$$\tilde{K} - \tilde{Q} + \tilde{P}_k = \tilde{P} + \tilde{\beta} \tag{3.68}$$

Applying the Johansen linearization to all of the PROTO equations results in the following linearized system:

[17] It could also be obtained by logarithmic differentiation. Johansen's method is identical to log differentiation for products but is slightly different for sums.

$$-\tilde{Q}+\tilde{X}=\tilde{\alpha} \tag{3.69}$$

$$\tilde{P}_k+\tilde{K}-\tilde{Q}-\tilde{P}=\tilde{\beta} \tag{3.70}$$

$$\tilde{L}-\tilde{Q}+\tilde{W}-\tilde{P}=-\frac{\alpha}{1-\alpha-\beta}\tilde{\alpha}-\frac{\beta}{1-\alpha-\beta}\tilde{\beta} \tag{3.71}$$

$$\begin{aligned}-\beta Q\tilde{K}-(1-\alpha-\beta)Q\tilde{L}+Q\tilde{Q}-\alpha Q\tilde{X}=\\ \alpha Q\ln\left(\frac{X}{L}\right)\tilde{\alpha}+\beta Q\ln\left(\frac{K}{L}\right)\tilde{\beta}+Q\tilde{A}\end{aligned} \tag{3.72}$$

$$PC(1+t)\tilde{C}-\gamma Y\tilde{\gamma}+PC(1+t)\tilde{P}+tPC\tilde{t}=\gamma Y\tilde{Y} \tag{3.73}$$

$$WJ\tilde{W}+WJ\tilde{J}-(1-\gamma)Y\tilde{Y}=-\gamma Y\tilde{Y} \tag{3.74}$$

$$-P_kK\tilde{K}-P_kK\tilde{P}_k-S\tilde{S}-WH\tilde{W}+Y\tilde{Y}-WH\tilde{H}=0 \tag{3.75}$$

$$J\tilde{J}+L\tilde{L}-H\tilde{H}=0 \tag{3.76}$$

$$\tilde{K}_s=0 \tag{3.77}$$

$$\tilde{C}-\tilde{S}+\tilde{P}+\tilde{t}=0 \tag{3.78}$$

The behavioral equations of the linearized model provide clear interpretations. For example, the linearized production function shows that the percent change in output will be equal to the weighted sum of percentage changes in all the inputs (in the absence of changes in the parameters). The weight associated with each input variable is the corresponding share parameter. Similarly, the consumption demand equation shows that the percentage change in consumption will be equal to the weighted difference in the percentage changes in full income and the output price (also in the absence of the change in the utility parameter). These weights are sometimes referred to as "S-elasticities."[18]

[18] Pagen, Adrian R., and John H. Shannon (1985)

The variable vectors for the linearized PROTO model, expressed in terms of percent changes, are shown below:

$$\tilde{m} = \left[\tilde{C}, \tilde{J}, \tilde{K}, \tilde{L}, \tilde{P}_k, \tilde{Q}, \tilde{S}, \tilde{W}, \tilde{X}, \tilde{Y}, \right]^T \tag{3.79}$$

$$\tilde{n} = \left[\tilde{H}, \tilde{P}, \tilde{t} \right]^T \tag{3.80}$$

$$\tilde{\theta} = \left[\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{A} \right]^T \tag{3.81}$$

The description of the endogenous, exogenous, and the parameter vectors are given in the tables below.

Table 3.6: Description of Endogenous Variables

Variables	Description
C	Consumption
J	Leisure
K	Capital
L	Labor
Pk	Rental price of capital
Q	Output
S	Subsidy
W	Wage rate
X	Intermediate good
Y	Full income

Table 3.7: Description of Exogenous Variables

Variables	Description
H	Total endowment of time
P	Output price
T	Tax rate on consumption

Table 3.8: Description of the Parameters

Variables	Description
α	Intermediate good's share parameter
β	Capital share parameter
γ	Utility parameter
A	Technology parameter

In the absence of any changes in the disturbance terms, the percentage change form of the model will be:

$$M\tilde{m} + N\tilde{n} = O\tilde{\theta} \tag{3.82}$$

The matrices M, N and O are shown in detail below.

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -\alpha Q & 0 & 0 & \alpha Q & 0 \\ 0 & 0 & P_k K & 0 & -P_k K & \beta P Q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\beta Q & -(1-\alpha-\beta)Q & 0 & Q & 0 & 0 & -\alpha Q & 0 \\ PC(1+t) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma Y \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -P_k K & 0 & -P_k K & -WH & -S & 0 & 0 & Y \\ 0 & J & 0 & L & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}$$

$$N = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\beta PQ & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & PC(1+t) & tPC \\ 0 & 0 & 0 \\ -WH & 0 & 0 \\ -H & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tag{3.83}$$

$$O = \begin{bmatrix} \alpha Q & 0 & 0 & 0 \\ 0 & \beta PQ & 0 & 0 \\ -\dfrac{\alpha}{1-\alpha-\beta} & -\dfrac{\beta}{1-\alpha-\beta} & 0 & 0 \\ \alpha Q \ln\left(\dfrac{X}{L}\right) & \beta Q \ln\left(\dfrac{K}{L}\right) & 0 & Q \\ 0 & 0 & \gamma Y & 0 \\ 0 & 0 & 0 & \dfrac{\gamma}{1-\gamma} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.84}$$

3.3.2 Model with Overall Uncertainty

In this section, overall confidence intervals for the model's endogenous variables are derived. The overall confidence interval includes both types of uncertainties: imprecision in the parameters and the residuals in the estimating equations. To begin the derivation, recall that the system of nonlinear equations can be written as:

$$F(m, n, \theta, \varepsilon) = 0 \tag{3.85}$$

The point prediction generated by the model will be a vector $\hat{m}$ that satisfies the following equation:

$$F(\hat{m}, n, \hat{\theta}, 0) = 0$$

where $\hat{\theta}$ is the vector of estimated parameters. The vector of exogenous variables, *n*, appears in both equations because it is predetermined and assumed to be known precisely. Constructing a Taylor Series expansion of the model around the base case solution for the true solution gives:

$$F_m(m - \hat{m}) + F_\theta(\theta - \hat{\theta}) + F_\varepsilon \Delta_\varepsilon = 0 \tag{3.86}$$

where F_m, F_θ and F_ε are components of the Jacobian matrix of F: matrices of partial derivatives of the equations with respect to each group of variables. To

convert this expression to its Johansen form, define matrices D_m and D_θ to be diagonal with the elements of $\hat{m}$ and $\hat{\theta}$, respectively, on their diagonals. Inserting D_m and D_θ:

$$F_m D_m D_m^{-1}(m-\hat{m}) + F_\theta D_\theta D_\theta^{-1}(\theta-\hat{\theta}) + F_\varepsilon \Delta_\varepsilon = 0$$

In terms of the matrices used in the previous section, $F_m D_m = M$, $-F_\theta D_\theta = O$ and $-F_\varepsilon = D$. Thus, the equation can be rewritten as shown:

$$MD_m^{-1}(m-\hat{m}) = OD_\theta^{-1}(\theta-\hat{\theta}) + D\Delta_\varepsilon \tag{3.87}$$

Solving for the prediction error in terms of the errors in the parameters and the residuals:

$$e_f = m-\hat{m} = D_m M^{-1} OD_\theta^{-1}(\theta-\hat{\theta}) + D_m M^{-1} D\Delta_\varepsilon \tag{3.88}$$

The variance of e_f will be:

$$\begin{aligned}\mathrm{var}(e_f) = &\\ &(D_m M^{-1} OD_\theta^{-1})\,\mathrm{var}(\hat{\theta})(D_m M^{-1} OD_\theta^{-1})^T + \\ &(D_m M^{-1} D)\,\mathrm{var}(\varepsilon)(D_m M^{-1} D)\end{aligned} \tag{3.89}$$

Alternatively, this can be expressed in terms of percentage changes from $\hat{m}$ and $\hat{\theta}$ as follows:

$$\text{var}(\tilde{e}) = (M^{-1}O)\,\text{var}(\hat{\tilde{\theta}})(M^{-1}O)^T + (M^{-1}D)\,\text{var}(\varepsilon)(M^{-1}D) \qquad (3.90)$$

Writing the variances in matrix notation allows this to be expressed as follows:

$$\tilde{\Sigma} = \left(M^{-1}\mathrm{O}\right)\underset{\tilde{\theta}}{\Sigma}\left(M^{-1}\mathrm{O}\right) + \left(M^{-1}D\right)\Sigma_{\varepsilon}\left(M^{-1}D\right) \qquad (3.91)$$

where the standard error of any given variable is the square root of the corresponding diagonal element:

$$\tilde{\sigma}_i = \sqrt{\tilde{\Sigma}_{ii}} \qquad (3.92)$$

and $\tilde{s}_i$ is an estimator for $\tilde{\sigma}_i$. If the error terms are normally distributed then the prediction error is sum of normal variables, and is thus normally distributed. Then, following the steps described in Chapter 2, the overall 95 percent confidence interval for the i^{th} endogenous variables, expressed in percent changes, can be written as:

$$\tilde{\hat{m}}_i + 2\tilde{s}_i \qquad (3.93)$$

Equation 3.77 gives the confidence interval in terms of percentage deviations in the model's variables relative to the base case. In addition, the interval can be expressed in terms of the absolute levels of the variables. That is, it could be written as:

$$\hat{m}_i + 2s_i \tag{3.94}$$

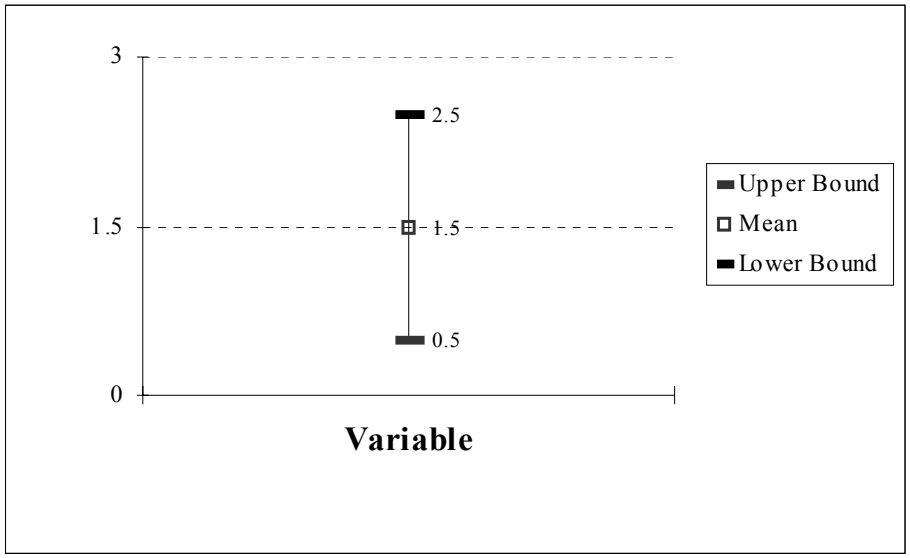
where s_i is obtained from equation 3.73, reproduced in matrix form below:

$$\Sigma = \left(\mathrm{D}_{\hat{m}} \mathrm{M}^{-1} O D_{\hat{\theta}}^{-1} \right) \Sigma_{\hat{\theta}} \left(\mathrm{D}_{\hat{m}} \mathrm{M}^{-1} O D_{\hat{\theta}}^{-1} \right)^T + \left(\mathrm{D}_{\hat{m}} \mathrm{M}^{-1} D^{-1} \right) \Sigma_{\varepsilon} \left(\mathrm{D}_{\hat{m}} \mathrm{M}^{-1} D^{-1} \right)^T \tag{3.95}$$

Equation 3.78 provides the general expression for overall confidence intervals that account for the variance-covariance matrix associated with the estimated parameters and for the unexplained variance in the regression equations. Equation 3.79 can be used to calculate s_i, the key ingredient in forming the confidence interval for variable i. As before, the prediction error can be decomposed into two components corresponding to the two major terms on the right side of equation 3.79: a portion that reflects the variance in the estimated parameters, and a portion that reflects the unexplained variance.

Conceptually, a confidence interval can be represented by a vertical line like the one as shown in Figure 3.3. The midpoint in the figure shows the mean of the hypothetical variable (1.5). The limits of the confidence interval are shown as the upper and lower bounds on the line segment (2.5 and 0.5; these values imply that the variable's standard error must be 0.5 if the diagram is displaying a 95 percent confidence interval). In terms of the percentage change from the mean, the width of the interval is approximately 66%.

Figure 3.3: Graphical Representation of a Confidence Interval



The above approach can now be applied to the full PROTO model. As mentioned previously, the disturbance terms are introduced into PROTO by the four econometric equations. The model with these equations inserted can be written as:

$$PX = \alpha PQ + \varepsilon_1 \tag{3.96}$$

$$P_K K = \beta PQ + \varepsilon_2 \tag{3.97}$$

$$\text{WL} = (1 - \alpha - \beta)\text{PQ} \tag{3.98}$$

$$Q = AX^{\alpha} K^{\beta} L^{1-\alpha-\beta} e^{\varepsilon_4} \tag{3.99}$$

$$PC(1+t) = \gamma Y + \varepsilon_3 \tag{3.100}$$

$$WJ = (1 - \gamma)Y \tag{3.101}$$

$$Y = WH + P_k K + S \tag{3.102}$$

$$H = L_s + J \tag{3.103}$$

$$K_s = \overline{K} \tag{3.104}$$

$$tPC = S \tag{3.105}$$

The linearized model differs only in the equations with the disturbance terms. As an illustration, the household demand for consumption goods could be linearized as follows. Expressed in levels, the equation has the form:

$$PC(1+t) = \gamma Y + \varepsilon_3$$

where ε_3 is the error term. Totally differentiating the equation with respect to the endogenous and exogenous variables, the parameter, and the residual gives:

$$d(PC(1+t)) = d(\gamma Y + \varepsilon_3) \tag{3.106}$$

Further expanding the above equation gives:

$$C(1+t)dP + P(1+t)dC + PCdt = \gamma dY + Yd\gamma + d\varepsilon_3 \tag{3.107}$$

The differences for the endogenous variables and the parameter are measured from the base case solution and the mean value of the estimate. The residual is linearized around its mean value, which is zero.[19] That is, the residual difference variable can be written as:

$$d\varepsilon_3 = \varepsilon_3 - \bar{\varepsilon}_3 = \varepsilon_3 \tag{3.108}$$

Equation 3.91 can be written in percentage change as:[20]

$$PC(1+t)\frac{dP}{P} + PC(1+t)\frac{dC}{C} + tPC\frac{dt}{t} = \gamma Y\frac{dY}{Y} + \gamma Y\frac{d\gamma}{\gamma} + d\varepsilon_3 \tag{3.109}$$

As before, this can be made more compact by using a tilde to represent the percentage change form of a variable:

---

[19] This form is chosen strictly for avoiding computation problems. The division by zero would blow up the program. This problem is not unique to the residuals only, it is also possible that one of the exogenous variables e.g., tax is zero in the base case solution but non-zero under a policy change. In such a situation, simulation requires that the tax variable be in other form than the percentage form. Besides the difference form it is also possible to redefine such variables so as to avoid this problem. See the GEMPACK program given in Appendix C.

[20] Note the possibility of a division-by-zero problem with the variable $\tilde{t}$ as explained in the previous footnote.

$$PC(1+t)\tilde{P}+PC(1+t)\tilde{C}+tPC\tilde{t}=\gamma Y\tilde{Y}+\gamma Y\tilde{\gamma}+d\varepsilon_3 \tag{3.110}$$

The complete list of linearized equations for the model, inclusive of the econometric equations, is shown below:

$$-\alpha Q\tilde{Q}+\alpha Q\tilde{X}=\alpha Q\tilde{\alpha}+d\varepsilon_1 \tag{3.111}$$

$$P_kK\tilde{P}_k+P_kK\tilde{K}-\beta PQ\tilde{Q}-\beta PQ\tilde{P}=\beta PQ\tilde{\beta}+d\varepsilon_2 \tag{3.112}$$

$$W\tilde{W}+L\tilde{L}-(1-\alpha-\beta)\left(P\tilde{P}+Q\tilde{Q}\right)=PQ\left(-\alpha\tilde{\alpha}-\beta\tilde{\beta}\right) \tag{3.113}$$

$$-\beta Q\tilde{K}-(1-\alpha-\beta)Q\tilde{L}+Q\tilde{Q}-\alpha Q\tilde{X}= \alpha Q\ln\left(\frac{X}{L}\right)\tilde{\alpha}+\beta Q\ln\left(\frac{K}{L}\right)\tilde{\beta}+Q\tilde{A}+d\varepsilon_4 \tag{3.114}$$

$$PC(1+t)\tilde{P}+PC(1+t)\tilde{C}+tPC\tilde{t}=\gamma Y\tilde{Y}+\gamma Y\tilde{\gamma}+d\varepsilon_3 \tag{3.115}$$

$$WJ\tilde{W}+WJ\tilde{J}-(1-\gamma)Y\tilde{Y}=-\gamma Y\tilde{\gamma} \tag{3.116}$$

$$-P_kK\tilde{K}-P_kK\tilde{P}_k-S\tilde{S}-WH\tilde{W}+Y\tilde{Y}-WH\tilde{H}=0 \tag{3.117}$$

$$J\tilde{J}+L\tilde{L}-H\tilde{H}=0 \tag{3.118}$$

$$\tilde{K}_s=0 \tag{3.119}$$

$$\tilde{C}-\tilde{S}+\tilde{P}+\tilde{t}=0 \tag{3.120}$$

In terms of matrix notation, the coefficient matrices *M*, *N*, *O* are the same as shown in the previous section. The coefficient matrix, *D* of the residual vector, $\Delta_\varepsilon=(d\varepsilon_1,d\varepsilon_2,d\varepsilon_3,d\varepsilon_4)$, can be written as:

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.121}$$

3.3.3 Model with Estimated Parameter Uncertainty

The confidence intervals for the mean values of the endogenous variables involve only the estimated parameter covariance. The procedure for calculating them is nearly identical to the one described in the above section. The prediction error for this case is given by:

$$e_f = m - \hat{m} = D_m M^{-1} O D_{\theta}^{-1} (\theta - \hat{\theta}) \tag{3.122}$$

Following steps outlined in the above section, the 95 percent confidence interval for the percentage deviation in $\hat{m}$ is

$$\tilde{\hat{m}}_i + 2\tilde{s}_i \tag{3.123}$$

where $\tilde{s}_i$ is the estimator of the i^{th} diagonal component of

$$\tilde{\Sigma} = \left(M^{-1}\mathrm{O}\right) \underset{\sim}{\Sigma}_{\theta} \left(M^{-1}\mathrm{O}\right)^{T} \tag{3.124}$$

As before, the confidence interval shown in equation 3.107 can also be expressed in terms of the levels. The principle difference between this confidence interval and the previous one is that component associated with the estimation residuals is absent. Thus, this interval communicates the sensitivity of the model to imprecision in the estimated parameters.

3.4 MODEL IMPLEMENTATION

In the above sections, an analytical approach for computing confidence intervals was described and the equations from the PROTO model were used to illustrate the process. The following sections build on this foundation to calculate numerical confidence intervals for PROTO.

The first step involved in the numerical implementation process is to solve for the model's base case solution. This is important because the variables in the linearized model are expressed as percentage changes from the base case solution. The second step in the process is to compute the key components of the model's Jacobian matrix (see equation 3.79 and 3.108). The final step is to use this information to calculate the actual confidence intervals. Thus, the whole process can be summarized as follows: linearizing the model, calculating the base case, evaluating the Jacobian matrix, and finally substituting all the information into the confidence interval expression.

The GEMPACK software package is used for several key steps in the computation. GEMPACK is a suite of general-purpose economic modeling programs that is especially suitable for general and partial equilibrium models. The software is very flexible and allows a model to be expressed algebraically in levels, differences, Johansen linearized form, or any combination of those forms. This eliminates the laborious task of manual linearization of large and complex models, and avoids the possibility of introducing errors due to incorrect manual linearization. The package includes multiple features and powerful linear algebra algorithms to ensure the accuracy of the solution. It is also extremely easy to use and very fast at performing simulations given a base case solution.

3.4.1 Computation of the Base Case Solution

Solving the base case requires solving a system of nonlinear equations. This could be done using other optimization software, e.g. GAMS, IML, CPLEX, etc., instead of GEMPACK. Using a separate package, however, would require programming in two different languages. For example, there would be one program for finding the base case solution and another for performing simulations in GEMPACK. It would be duplicative modeling work at best, and at worst it would be a vast opportunity for programming errors to be introduced. Furthermore, an in-depth knowledge of both packages would be required because the tasks are computationally challenging. In this section, an approach to computing the base case solution is explained that avoids the need for a second

software package. It can be easily implemented in GEMPACK and draws on work done by Codsi, et al., (1992) and Horridge, et al., (1993).

The first step is to append a slack variable to each of the equations of the original model (3.44). The new system of equations with the slack variables will be referred to as the "slack-augmented system" and can be written as shown:

$$F(m, n, \theta) + S = 0 \tag{3.125}$$

where S is p x 1 vector of slack variables.

The equations are then used to compute the initial value of each slack variable by calculating how far apart the left and right hand sides of each equation would be under the initial dataset and in the absence of its slack variable. The resulting values of the slack variables complete the slack-augmented system by causing 3.109 to hold.[21] For example, if m_0 is a vector of initial starting guesses for the endogenous variables, then a corresponding slack variable S_0 would be calculated as shown:

$$S_0 = -F(m_0, n, \theta) \tag{3.126}$$

[21]"Slack augmented system," represents reality. It is not always the case that equations of the economy are free from errors. These errors are due to discrepancies in calculating the economy identity equation. So, in essence, if one was to start the model with the existing data, then the slacks could be identified as the error in data. It is possible to carry along these slacks with model analysis. Solution for this model satisfies all the equations without any slack values.

The next step is then to linearize the slack-augmented system. In matrix form the Johansen linearized model can be written as:

$$M\tilde{m} + N\tilde{n} = O\tilde{\theta} + I\tilde{S} \tag{3.127}$$

where $\tilde{m}, \tilde{n}, \tilde{\theta}$ and $\tilde{S}$ are the endogenous and exogenous variables, the vector of parameters, and the vector of slack variables, all expressed in percentage change form. I is a p x p identity matrix. All the other notation is similar to the previous sections.

A simulation is then performed on the slack-augmented system whereby the slacks are reduced by 100 percent. That is, $I\tilde{S}$ is replaced by a p x p diagonal matrix with negative 100 on its diagonal elements and the exogenous variables and parameters are kept at their base case values. By performing this simulation, the above equation 3.111 reduces to:

$$M\tilde{m} = I_{-100} \tag{3.128}$$

Equation 3.112 is then solved for new values of the endogenous variables. The results are used to update the initial guess of m as follows:

$$m_1 = m_0 + D_{m_0} M^{-1} I_{-1} \tag{3.129}$$

Performing this experiment drives the slack variables from their initial values (S_0) to zero. The result is a solution to the original model. However, the solution constructed at this point is only an approximation to the nonlinear solution because it is obtained by solving the linearized equations of the model only once. The updated solution from this experiment, m_1, is referred to as "single step Johansen solution" to the model. The base case solution can be refined much further, and brought arbitrarily close to the true nonlinear solution, by repeating the process several times in a "multi-step Johansen solution": calculating new slack values, S_1, for the post-simulation version of the model; updating the base case data set; and then repeating the simulation shocking the slack variables to zero. The number of iterations required to push the slack variables close to zero depends upon the curvature of the functions in the model, and on the quality of the initial guess.

Performing multi-step Johansen simulations, using Euler's method or other solution algorithms available in GEMPACK, can enhance the accuracy of the solution.[22] A detailed explanation of the GEMPACK programs for evaluating the base case solution for PROTO is presented in the Appendix C.

3.4.2 Computing Matrices Derived from the Jacobian

Given the base case, the next step is to compute the key matrices needed for constructing confidence intervals. Each of these matrices is linked to a portion of the Jacobian matrix for the model. GEMPACK does not include a routine

[22] See GEMPACK manual for detailed explanation of multi-step simulation in order to get accurate solutions of nonlinear equations.

specifically designed for computing these matrices directly; however, it is possible to calculate them by performing appropriate simulations.

As an illustration of the procedure, consider constructing the matrix $M^{-1}O$. Rearranging equation 3.44 (without the error term) in order to solve for percentage changes in the model's endogenous variables:

$$\tilde{m} = \left(M^{-1}O\right)\tilde{\theta}_1 \tag{3.130}$$

One column of the desired matrix can be obtained by solving the above system when $\tilde{\theta}_1$ is a unit vector. That is, a simulation is performed where the percentage change in a single one of the parameters, $\tilde{\theta}_1$, is one percent. The resulting $\tilde{m}$ will be the corresponding column of $M^{-1}O$. A series of such experiments can be used to extract all of the columns. Moreover, GEMPACK has built-in capabilities for running large groups of simulations simultaneously and efficiently.

In passing, note that the *ij* element of $M^{-1}O$ is the percentage change in endogenous variable *i* due to a unit percentage change in parameter *j*. The weighted sum of all the percentage changes along row *i* gives the percentage change in variable *i* produced by a specified set of percentage changes in the parameters. This additivity is a powerful tool to compute effects on the endogenous variables for selective parameter change.[23]

Once the key matrices have been built using GEMPACK, the actual confidence intervals can be calculated by straightforward linear algebra. For

[23] This is also known as superposition principle.

convenience, the computations were done using the Gauss programming language. The relevant Gauss program is attached in Appendix B. The following section shifts the discussion from theory to practice and presents a detailed numerical analysis of the confidence intervals for the PROTO model.

3.5 RESULTS

3.5.1 Base Case Solution

The base case is solved by the iterative procedure described above. The year examined was chosen to be 1982. The output value amounts to $4,295 billion dollars. The consumer supplies 1.25 million units of labor out of the total time endowment of 8.9 million units, and it consumes $2,196 billion dollars worth of aggregate consumption good. The full income of the consumer was $11,182 billion dollars. The complete solution to the PROTO model is shown in Table 3.9.

Table 3.9: Base Case Solution (values in dollars)

Variables	Values	Units	Variables	Values	Units
C	$2196	mil.	Q	$4295	bil.
J	7.659	mil. hours	S	$219	bil.
K	928	bil. units	W	$1.144	dollars
L	1.249	mil. hours	X	$2099	bil.
P_k	$0.824	dollars	Y	$11182	bil.

3.5.2 Confidence Intervals with Parameter Uncertainty

This section presents confidence intervals for the PROTO model's endogenous variables, taking parameter uncertainty into account. In addition, confidence intervals are compared and contrasted with sensitivity analysis results.

Confidence intervals expressed in absolute levels (as opposed to percentage deviations from the base case) are shown for all of the endogenous variables in Figures 3.4a - 3.4i. The results are also given as percentage and absolute deviations from the base case solution in Table 3.10 and shown as percentages in Figure 3.5.

The confidence interval for consumption, for example, is about 4.7 percent and, in terms of levels, extends from $2,092 to $2,297 billion. The narrowest interval in percentage terms is about 2.5 percent for full income; the widest interval is 9 percent for labor supply. Note that in percentage terms, the confidence interval for labor supply is much larger than that for its counterpart variable, leisure. The reason is that labor supply is only about one eight of the total time endowment. A given change in labor supply, therefore, corresponds to a much smaller percentage change in leisure. In absolute terms, the confidence intervals for labor and leisure are identical.

Figure 3.4: Confidence Interval for Some Key Variables

Figure 3.4a: Confidence Interval for Consumption (bil. $)

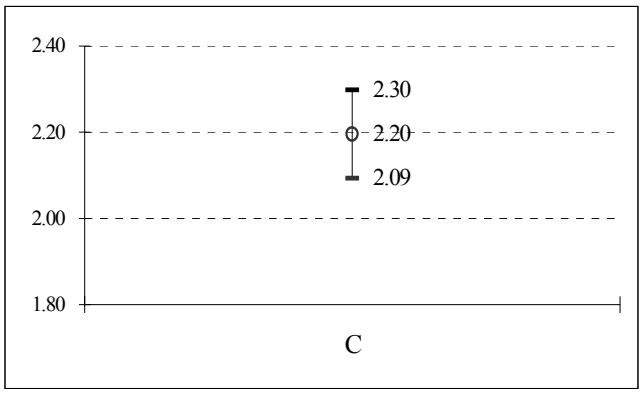


Figure 3.4b: Confidence Interval for Leisure (mil. hours)

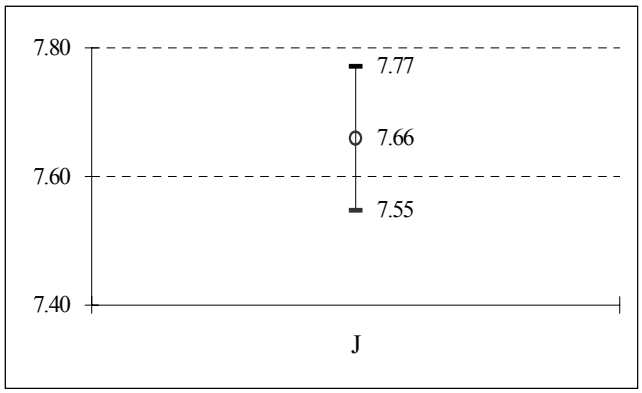


Figure 3.4c: Confidence Interval for Labor (mil. hours)

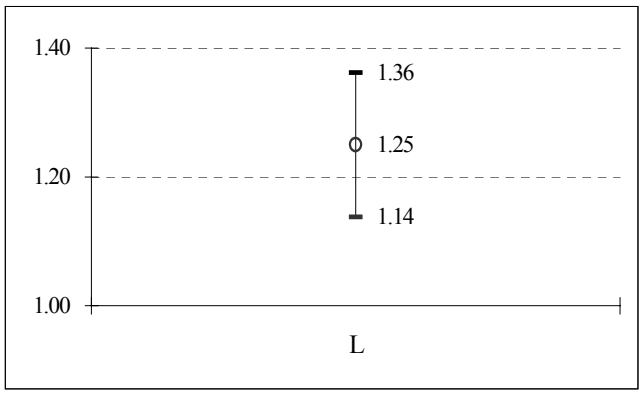


Figure 3.4d: Confidence Interval for Rental Price of Capital ($)

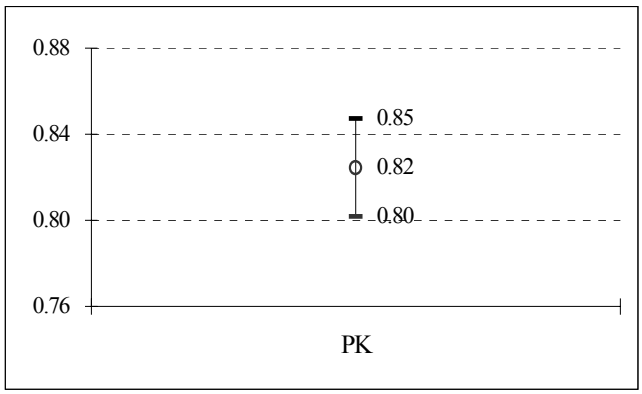


Figure 3.4e: Confidence Interval for Output (tril. $)

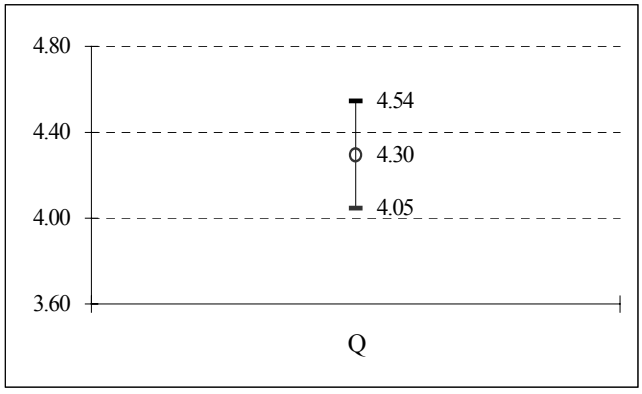


Figure 3.4f: Confidence Interval for Subsidy (tril. $)

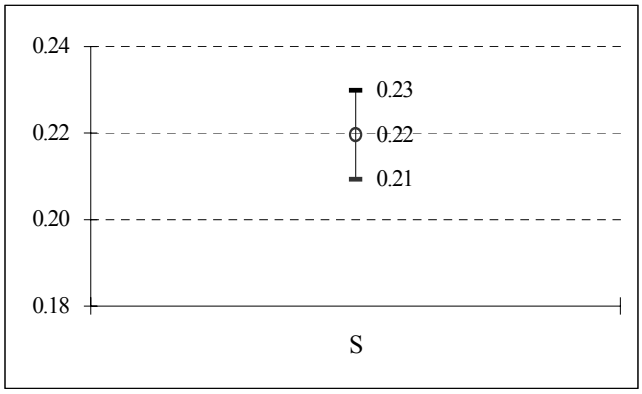


Figure 3.4g: Confidence Interval for Wage Rate ($)

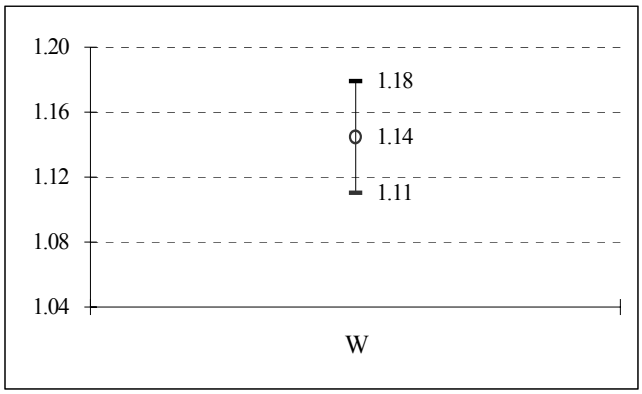


Figure 3.4h Confidence Interval for Intermediate Goods (tril. $)

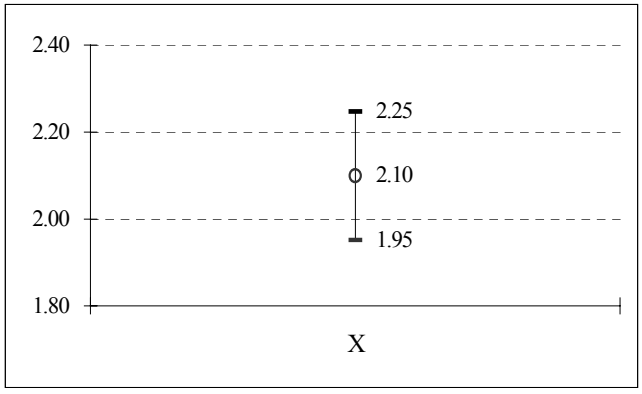


Figure 3.4i: Confidence Interval for Full Income (tril. $)

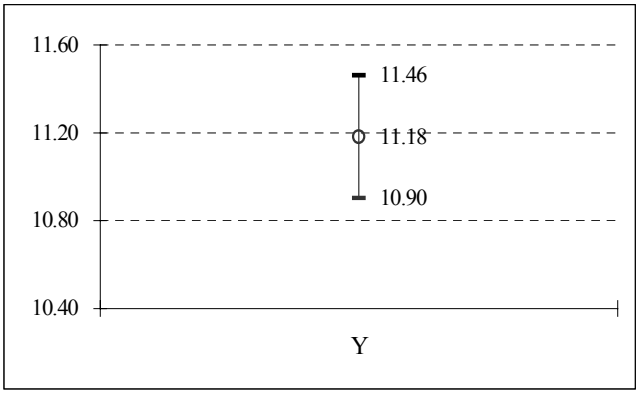
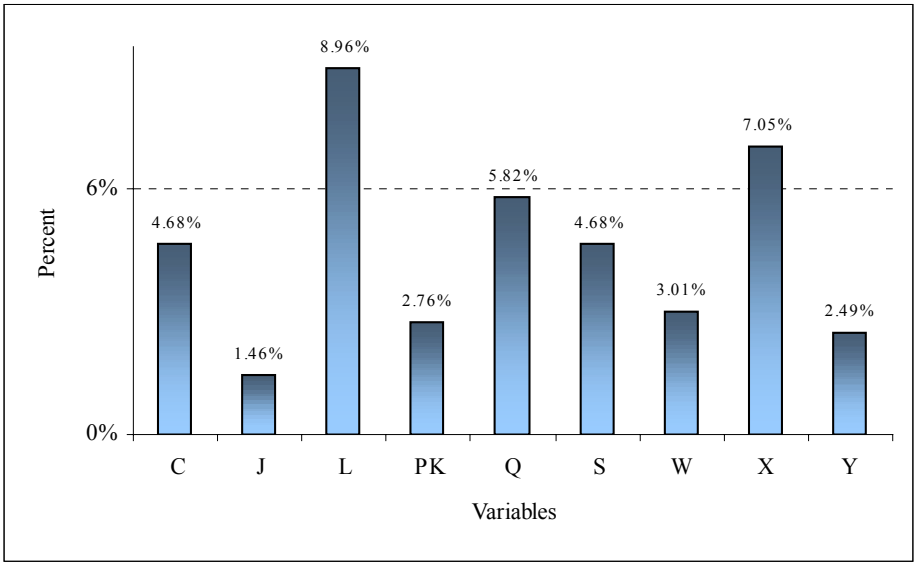


Table 3.10: Confidence Intervals as Percentage Change from the Base case Solution and the Standard Errors

Variables	Percent Deviation	2*Standard Errors
C	4.68%	$102,735
J	1.46%	$112,039
L	8.96%	112,039
Pk	2.76%	$ 0.0227
Q	5.82%	$249,850
S	4.68%	$10,274
W	3.01%	$0.0344
X	7.05%	$148,023
Y	2.49%	$278,603

Figure 3.5: Confidence Intervals in Percent Deviation from the Base Case Solution



Since PROTO is a small model, it is feasible to use a Monte-Carlo simulation to check the accuracy of the confidence intervals computed above.[24] An essential feature of the Monte-Carlo process is drawing a large random sample from the joint distribution of the parameters and residuals. The section below explains how the sample was drawn. It begins with a simple univariate case first and then moves on to the multivariate case.

The first step in the procedure for generating random samples from a univariate normal distribution, $N(\mu,\sigma^2)$, involves drawing random samples, *Y*, from a standard normal distribution, $N(0,1)$. These are then fed into the transformation function: $X = \mu + cY$. If $c = \sigma$ then the transformation produces random samples for the desired distribution $N(\mu,\sigma^2)$.

The multivariate case is similar. Let $X = (X_1,.....,X_p)^T$ denote a random vector with multivariate normal probability distribution function:[25]

$$f(x) = (2\pi)^{-p/2}|\Sigma|^{-1/2} e^{-(x-\mu)^T \Sigma^{-1}(x-\mu)} \qquad -\infty < x_i < \infty \quad i = 1,...,p \quad (3.131)$$

where: μ is *p* x *1* column vector of means (μ_1,..., μ_p) and Σ denotes an *r* x *r* positive definite, symmetric variance-covariance matrix. The expression in 3.115 can be denoted compactly by $N(\mu,\Sigma)$. Further, let $Y = (Y_1,...,Y_r)^T$ be a sample drawn from a distribution of independent, standardized normal variables, ***N(0,I)***,

[24] The Jacobian matrix for the confidence interval calculation is derived from a single step simulation. Thus results from Monte Carlo and confidence intervals modeling could very likely be different.

[25] For detailed explanation on sample generating algorithms see Fishman (1996).

where $\boldsymbol{0}$ is a column vector of r zeros and $\boldsymbol{I}$ is the r x r identity matrix. Then a transformation can be written as:

$$X = \mu + cY \tag{3.132}$$

where $\boldsymbol{c}$ is a nonsingular r x r matrix satisfying:

$$\Sigma = cc^T . \tag{3.133}$$

The matrix $\boldsymbol{c}$ is computed by transforming the covariance matrix using its eigenvalues and eigenvectors. The symmetric positive definite matrix Σ can be diagonalized by the orthogonal matrix of eigenvectors, that is,[26]

$$\Xi^T \Sigma \Xi = \Lambda \tag{3.134}$$

where: Ξ is the orthogonal matrix of eigenvectors and Λ is diagonal matrix containing the eigenvalues (λ) of Σ in descending order on the principal diagonal.

The eigenvalues of the covariance matrix are all positive because it is positive definite. Thus the diagonal matrix can be factored into two matrices. Rearranging equation 3.118 gives:[27]

---

[26] Johnston (1984), pp. 147.

[27] Note that the property to transpose an orthogonal matrix is equal to the inverse of the orthogonal matrix is also used to simplify the expression, $\Xi^T = \Xi^{-1}$.

$$\Sigma = \Xi\Lambda^{1/2}\Lambda^{1/2}\Xi^{T} = \left(\Xi\Lambda^{1/2}\right)\left(\Xi\Lambda^{1/2}\right)^{T} \tag{3.135}$$

where: $\Lambda = diag\{\lambda_1,...,\lambda_r\}$.

Comparing equation 3.119 with equation 3.117 gives the ***c*** matrix as:

$$c = \left(\Xi\Lambda^{1/2}\right) \tag{3.136}$$

Substituting ***c*** from equation 3.120 into equation 3.116 gives the expression for generating random samples from a joint normal distribution using the transformed expression shown below:[28]

$$X = \mu + \left(\Xi\Lambda^{1/2}\right)Y \tag{3.137}$$

For PROTO, each simulation of the Monte-Carlo analysis involves drawing a *4* x *1* vector from the distribution of parameter estimates. This vector is then multiplied by the matrix ***c***, which is computed using the eigenvectors and eigenvalues of the parameter variance-covariance matrix. This resulting vector constitutes the random component that is added to the mean values of the parameters (μ). This yields one draw of the parameter vector used in the Monte-

[28] A straightforward application of following expectations of the expression 3.126 gives $E(X) = E\left(\mu + \Xi\Lambda^{1/2}Y\right) = \mu$ and $E\left[(X-\mu)(X-\mu)^{T}\right] = \Xi\Lambda^{1/2}\Lambda^{1/2}\Xi^{T} = \Sigma$.

Carlo simulation. The total number of parameter vectors drawn was 10,000 and they were constructed using a Gauss program.

The Monte-Carlo results are shown as a series of panels in Figure 3.6. Each shows the distribution of one of the endogenous variables. For example, Figure 3.6a shows the distribution of consumption. The 95% confidence interval calculated from the above runs from 2.13 to 2.41 trillion dollars and falls within the appropriate region of the distribution. The 95% confidence interval of leisure, which runs from 7.73 to 7.43 million hours, is also appropriate.[29] In fact, all of the confidence intervals calculated above are consistent with the corresponding Monte-Carlo results. The proposed approach for computing CGE confidence intervals generates good approximations to the true confidence intervals for the full nonlinear version of PROTO.

[29] 95% confidence intervals for the Monte-Carlo simulations were approximated by taking the 251^{th} and the 9750^{th} values on the sorted data result.

Figure 3.6: Probability Distribution for Some Key Variables

Figure 3.6a: Probability Distribution of Consumption (tril.$)

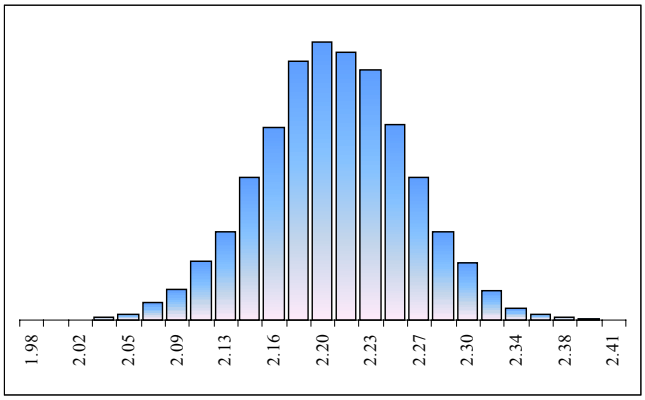


Figure 3.6b: Probability Distribution of Leisure (mil. hours)

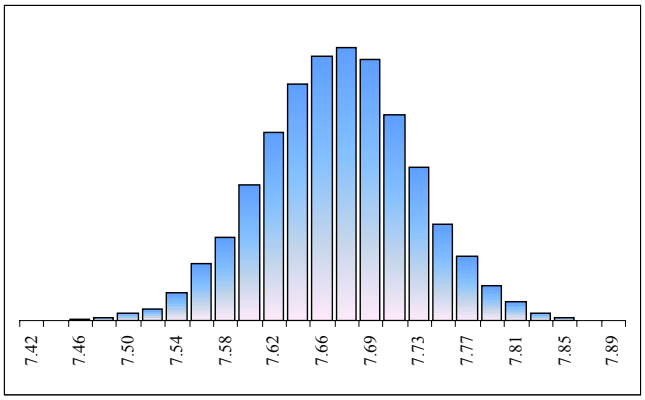


Figure 3.6c: Probability Distribution of Labor (mil. hours)

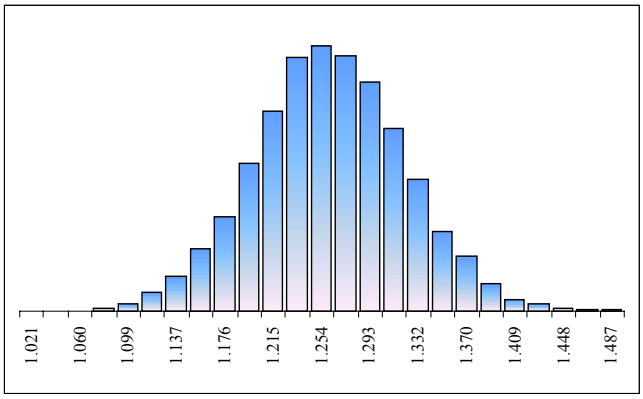


Figure 3.6d: Probability Distribution of Rental Price of Capital ($)

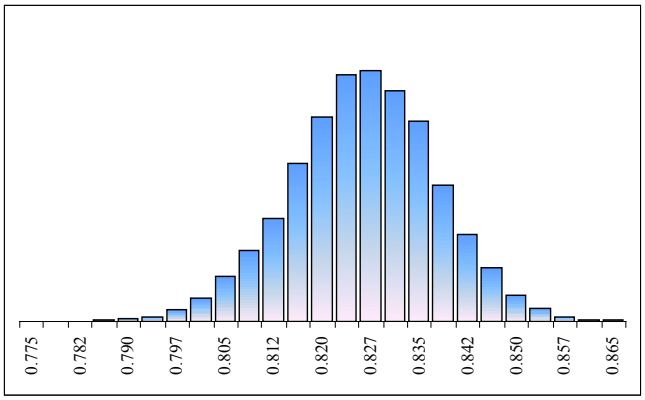


Figure 3.6e: Probability Distribution of Output (tril. $)

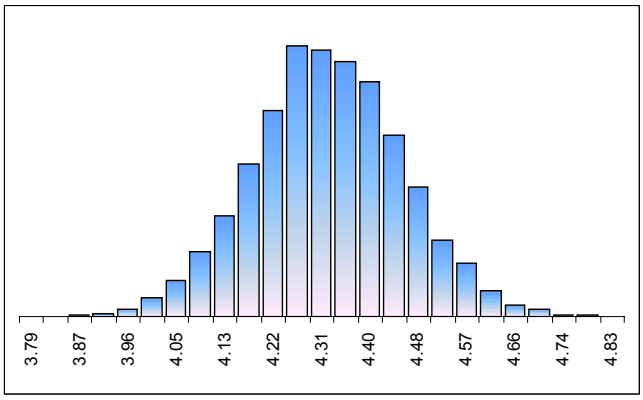


Figure 3.6f: Probability Distribution of Subsidy (tril. $)

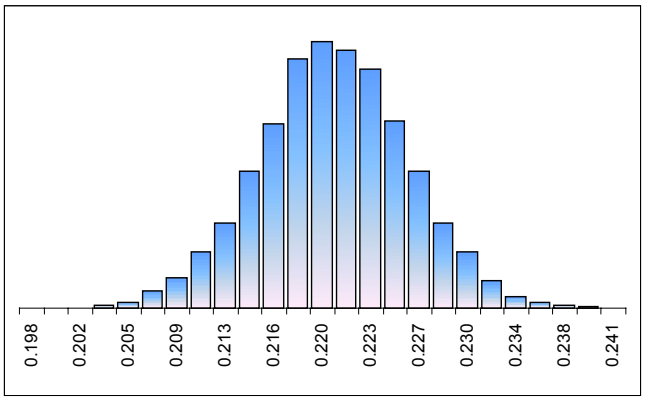


Figure 3.6g: Probability Distribution of Wage Rate ($)

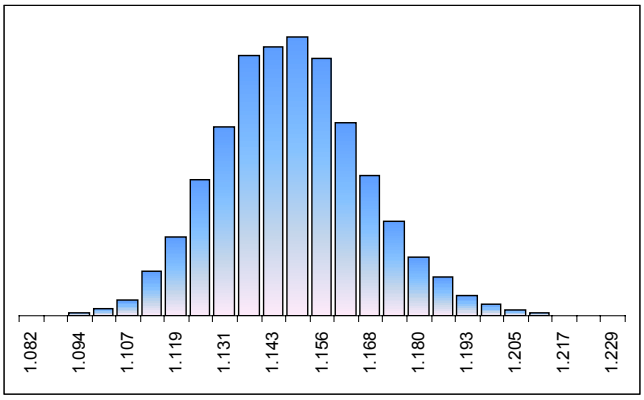


Figure 3.6h: Probability Distribution of Intermediate Goods (tril. $)

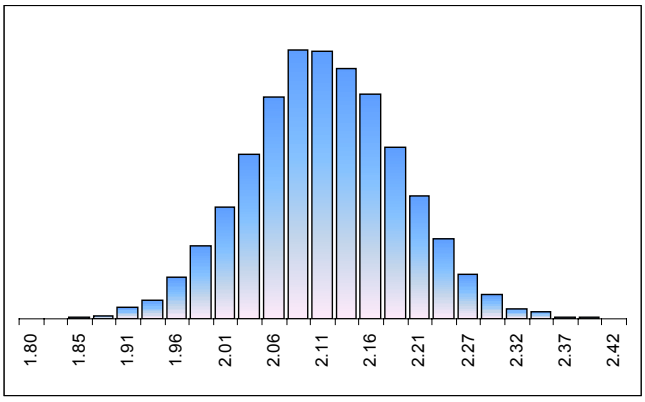
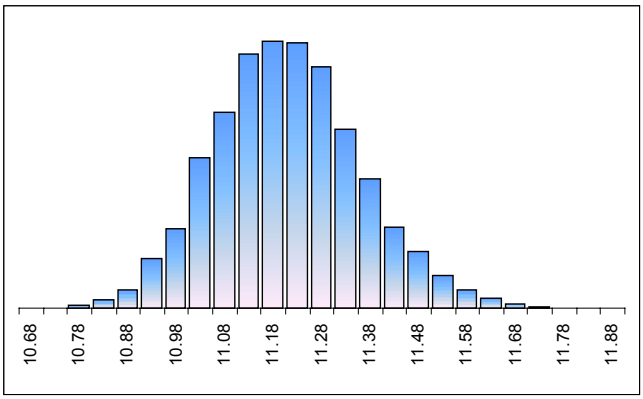


Figure 3.6i: Probability Distribution of Full Income (tril. $)



3.5.3 Confidence Intervals with Residual Uncertainty

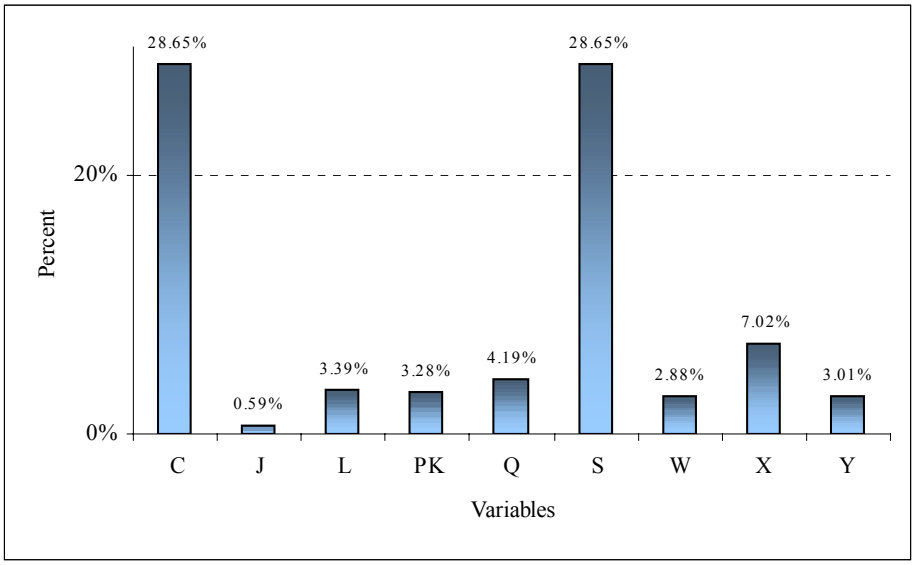
It is interesting to compute an alternative measure of the model's precision: the confidence intervals for the projected levels of the endogenous variables accounting for only the residuals in the estimation equations. Doing so provides a rough view of the asymptotic confidence intervals that would apply if the model were to be estimated using a very large data set. In effect, it allows the overall explanatory power of the estimating equations—or the lack thereof—to be mapped into the precision of model's endogenous variables. The larger the unexplained variance in the estimating equations, the wider will be the confidence intervals.

The results are shown in Table 3.11, which gives the confidence intervals due to unexplained variance as percentage changes from the corresponding variable's base case solution. The intervals vary widely across the variables, ranging from about 0.6 to 28 percent. Consumption and the subsidy are very imprecise: they each have confidence intervals equal to about 28 percent of their base case values. However, the confidence intervals for the other variables are considerably smaller.

Table 3.11: Confidence Intervals as Percentage Change from the Base Case Solution and Standard Errors Accounting for Residual Uncertainty

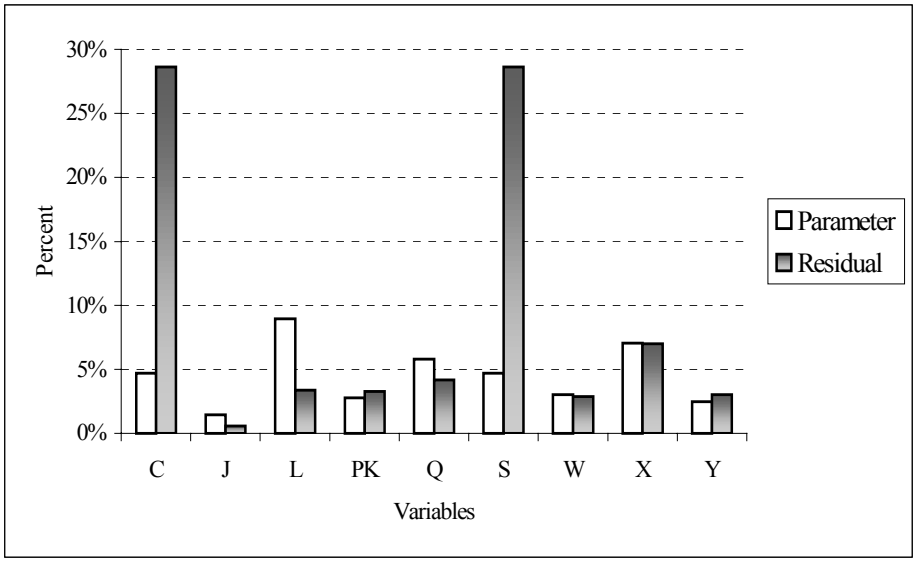
Variables	Percent Deviation	2*Standard Errors
C	28.65%	$629,004
J	0.59%	$45,436
L	3.39%	42,348
Pk	3.28%	$0.0270
Q	4.19%	$180,127
S	28.65%	$62,900
W	2.88%	$0.0330
X	7.02%	$147,326
Y	3.01%	$336,029

Figure 3.7: Confidence Intervals in Percent Deviation from the Base Case Solution Accounting for Residual Uncertainty



The confidence intervals for parameter and residual uncertainty are compared in the figure below. Consumption and the subsidy have confidence intervals due to residual uncertainty that are about six times larger than the confidence intervals due to parameter uncertainty. However, the confidence intervals stemming from parameter uncertainty are larger than those for the residuals for a number of variables, including leisure, output and the wage rate.

Figure 3.8: Confidence Intervals in Percent Deviation from the Base Case Solution



3.5.4 Numerical Results for the Overall Confidence Intervals

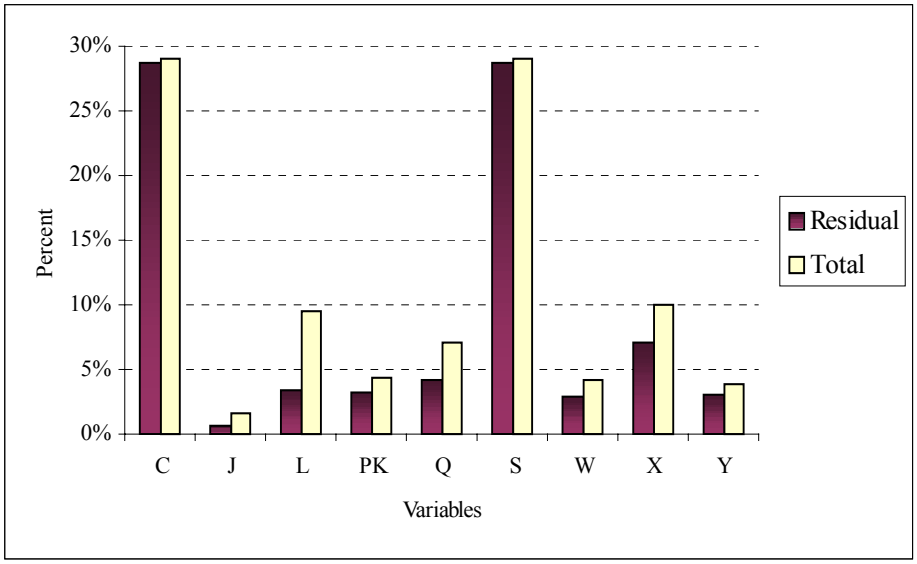
The overall confidence intervals, accounting for both uncertainty in the parameter estimates and the residuals in the estimating equations, are presented in percentage form in Table 3.12. Figure 3.9 shows the results graphically and

illustrates the relationship between the overall confidence intervals and those due only to uncertainty in the residuals.

Table 3.12: Confidence Intervals in Percentage Change Accounting for Overall Uncertainty

Variables	Percent Deviation
C	29.03%
J	1.58%
L	9.58%
Pk	4.28%
Q	7.17%
S	29.03%
W	4.16%
X	9.95%
Y	3.90%

Figure 3.9: Overall Confidence Intervals in Percentage Change and Confidence Interval Accounting for Unexplained Variation



3.6 COMPARISON WITH SENSITIVITY ANALYSIS

The analysis so far has focused on calculating confidence intervals. However, most modelers who have examined the precision of their models have used sensitivity analysis instead. Two main types of sensitivity analysis, described below, have been used.[30] This section presents both methods, applies them to PROTO, and compares the results with the confidence intervals reported above.

The first approach to sensitivity analysis entails perturbing the values of selected parameters to see how the model's results change. This approach will be referred as "traditional sensitivity analysis". Although this approach is simple to perform, it has several important drawbacks. First, there is no rigorous method for choosing which parameters to perturb. Second, there may be little basis for choosing the magnitude of each perturbation. In the absence of the statistical information provided by estimation, a modeler may have little alternative to picking arbitrary perturbations, such as halving and doubling the parameter. This amounts to imposing an ad hoc variance.[31]

The second approach to sensitivity analysis was suggested by Harrison, et al., (1993). It proceeds by first constructing a very coarse discrete approximation to the distribution of each parameter. Then simulations are run to determine the values of the endogenous variables for all possible combinations of the

[30] This section draws on Harrison, et al., (1993).

[31] Such self-imposed practice makes questionable the validity of taking all the graduate econometric courses!

parameters.[32] This approach is referred to as "unconditional systematic sensitivity analysis" (USSA). The main disadvantage with this approach is that the number of simulations required increases rapidly with the complexity of the model and the level of detail in the distribution of each parameter. For example, it might use five values for each of the four parameters of the PROTO model, in which case the number of simulations performed under USSA is 625. However, if the number of values used for each parameter were increased to ten, then the number of simulations performed would jump to 10,000. Increasing the number of parameters is even worse: using only three values for each of ten parameters would require nearly 60,000 simulations. The large number of simulations required by this approach gives it little or no significant advantage over Monte-Carlo simulation. Also, the large number of simulation results would be challenging to summarize concisely.

One way to reduce the number of parameter sets is to use “conditional systematic sensitivity analysis” (CSSA) by perturbing the parameter values selectively. With CSSA, a single parameter is perturbed while the rest of the parameters are held at their estimated values. This approach is subject to the same drawbacks as traditional sensitivity analysis: no rigorous basis for choosing which variables to perturb, or for deciding how much to perturb them.

For this exercise, three different parameter values are considered: "Most-Likely," "Likely," and "Unlikely." For the purpose of comparing the results with confidence intervals, the parameter values are evaluated at the mean, one standard

[32] The discrete approximation of the parameter values depends upon the assumed distribution of the parameters. Harrison, Vinod (1989) suggest that any on of the three distributions: uniform, t, and normal, can be assumed and sufficient in practice.

error, and two standard errors of the parameters (see Table 3.13).[33] Note that the parameter values do not vary substantially. This is because the parameter estimates are precisely determined and have small variances.

Table 3.13: Parameter Values for Sensitivity Analysis

Parameter	"Most Likely"	"Likely"	"Unlikely"
α	0.4888	0.4920	0.4952
β	0.1781	0.1817	0.1852
γ	0.2160	0.2237	0.2315
A	2.8123	2.8288	2.8453

Under CSSA, using three values for each of the four parameters in PROTO requires only twelve simulations. This approach is considerably less burdensome than USSA but it is unable to capture any covariance among the parameters. The results are shown in the tables to follow. Table 3.14 presents the results when the α parameter is increased by one and two standard errors (eg, to the "Likely" and "Unlikely" values in the table above). Under the "Unlikely" scenario, a change in the α parameter results in a 2.9 percent increase in the use of the aggregate intermediate good and a 0.6 percent decrease in the use of labor. This is a direct result in increasing the share of intermediate good in production. The deviations of the variables range from -0.59 to 2.9 percent.

Similarly, increasing the β parameter, as shown in Table 3.15, causes almost all of the variables decrease except the rental price of capital and the

[33] This approach of selecting the values of the parameters for conducting sensitivity analysis is more rationally correct than the procedure where it is arbitrarily selected.

quantity of leisure. The deviations in this case range from negative 1.8 percent to positive 2.3 percent. The fixed capital stock and an increase in the β parameter lead to an increase in the rental price. Wages fall and labor supply decreases as a result. Consumption is lower as well.

The γ parameter clearly has the most significant impact on the values of the endogenous variables, as shown by Table 3.16. In particular, labor supply increases by 7.9 percent, while consumption, the rental price of capital, total output, the subsidy, and the demand for intermediate goods, change by about 5.1 percent.[34]

Results for the technology parameter look sharply different, as shown in Table 3.17. All deviations are positive, reflecting the fact that an increase in A shifts the economy's production possibility frontier outward. An interesting observation for this table, however, is that all the variables that deviate from the base case solution do so by equal percentages. Again, this reflects the uniform outward shift in the production possibility frontier.

Further results for sixteen other simulations, in which multiple parameters are simultaneously perturbed by two standard errors, are shown in Table 3.18. For example, the simulation "0022" holds α and β at their estimated values and perturbs γ and A by two standard errors. It generates the largest impact on the endogenous variables. However, simulation "0002" has much less effect on the variables than "0022". This leads to the conclusion that γ is likely to be the most critical parameter for the model's results.

[34] An additional eight simulations were conducted in which the parameters were decreased by one and two standard deviations. The results showed the identical pattern observed with positive deviations.

Table 3.14: Model Solutions in Levels and Percentage Change from Base Case Solution Varying $\hat{\alpha}$

Variables	Likely		Unlikely	
	Levels	Deviation	Levels	Deviation
C	2,198,928	0.143%	2,202,305	0.297%
J	7,662,744	0.047%	7,666,430	0.096%
L	1,246,256	-0.291%	1,242,570	-0.586%
P_K	0.830895	0.776%	0.837468	1.574%
Q	4,328,372	0.776%	4,362,612	1.574%
S	219,893	0.143%	220,230	0.296%
W	1.1457	0.096%	1.1469	0.201%
X	2,129,444	1.439%	2,160,307	2.909%
Y	11,198,150	0.143%	11,215,340	0.297%

Table 3.15: Model Solutions in Levels and Percentage Change from Base Case Solution Varying $\hat{\beta}$

Variables	Likely		Unlikely	
	Levels	Deviation	Levels	Deviation
C	2,178,464	-0.789%	2,161,394	-1.566%
J	7,670,448	0.148%	7,681,822	0.297%
L	1,238,552	-0.907%	1,227,178	-1.817%
P_K	0.834108	1.166%	0.843565	2.313%
Q	4,261,138	-0.789%	4,227,750	-1.566%
S	217,846	-0.789%	216,139	-1.567%
W	1.1339	-0.936%	1.1234	-1.857%
X	2,082,674	-0.789%	2,066,355	-1.566%
Y	11,093,930	-0.789%	11,007,000	-1.566%

Table 3.16: Model Solutions in Levels and Percentage Change from Base Case Solution Varying $\hat{\gamma}$

Variables	Likely		Unlikely	
	Levels	Deviation	Levels	Deviation
C	2,251,813	2.551%	2,307,485	5.087%
J	7,609,830	-0.643%	7,560,208	-1.291%
L	1,299,172	3.943%	1,348,793	7.913%
P_K	0.845531	2.552%	0.866435	5.087%
Q	4,404,610	2.551%	4,513,506	5.087%
S	225,181	2.551%	230,748	5.087%
W	1.1293	-1.338%	1.1147	-2.618%
X	2,152,797	2.551%	2,206,022	5.087%
Y	11,070,800	-0.996%	10,965,210	-1.940%

Table 3.17: Model Solutions in Levels and Percentage Change from Base Case Solution Varying $\hat{A}$

Variables	Likely		Unlikely	
	Levels	Deviation	Levels	Deviation
C	2,221,005	1.148%	2,246,361	2.303%
J	7,659,106	0.000%	7,659,106	0.000%
L	1,249,893	0.000%	1,249,894	0.000%
P_K	0.833962	1.148%	0.843484	2.303%
Q	4,344,347	1.148%	4,393,948	2.303%
S	222,101	1.149%	224,636	2.303%
W	1.1578	1.149%	1.1710	2.303%
X	2,123,343	1.148%	2,147,586	2.303%
Y	11,310,570	1.148%	11,439,700	2.303%

Table 3.18: Parameter Values for Sensitivity Analysis

Parm.	Sim 2222	Sim 2220	Sim 2202	Sim 2200	Sim 2022	Sim 2020	Sim 2002	Sim 2000	Sim 0222	Sim 0220	Sim 0202	Sim 0200	Sim 0022	Sim 0020	Sim 0002
α	0.4952	0.4952	0.4952	0.4952	0.4952	0.4952	0.4952	0.4952	0.4888	0.4888	0.4888	0.4888	0.4888	0.4888	0.4888
β	0.1852	0.1852	0.1852	0.1852	0.1781	0.1781	0.1781	0.1781	0.1852	0.1852	0.1852	0.1852	0.1781	0.1781	0.1781
γ	0.2315	0.2315	0.2160	0.2160	0.2315	0.2315	0.2160	0.2160	0.2315	0.2315	0.2160	0.2160	0.2315	0.2315	0.2160
A	2.8453	2.8123	2.8453	2.8123	2.8453	2.8123	2.8453	2.8123	2.8453	2.8123	2.8453	2.8123	2.8453	2.8123	2.8453

Sensitivity Analysis Results in Percent Change

Var.	Sim 2222	Sim 2220	Sim 2202	Sim 2200	Sim 2022	Sim 2020	Sim 2002	Sim 2000	Sim 0222	Sim 0220	Sim 0202	Sim 0200	Sim 0022	Sim 0020	Sim 0002
C	6.03%	3.61%	1.02%	-1.29%	7.83%	5.37%	2.64%	0.30%	5.73%	3.35%	0.70%	-1.57%	7.51%	5.09%	2.30%
J	-0.87%	-0.87%	0.40%	0.40%	-1.19%	-1.19%	0.10%	0.10%	-0.98%	-0.98%	0.30%	0.30%	-1.29%	-1.29%	0.00%
L	5.32%	5.32%	-2.43%	-2.43%	7.29%	7.29%	-0.59%	-0.59%	5.98%	5.98%	-1.82%	-1.82%	7.91%	7.91%	0.00%
PK	11.61%	9.06%	6.33%	3.91%	9.20%	6.71%	3.94%	1.57%	9.90%	7.42%	4.67%	2.31%	7.51%	5.09%	2.30%
Q	7.38%	4.93%	2.30%	-0.03%	9.20%	6.71%	3.94%	1.57%	5.73%	3.35%	0.70%	-1.57%	7.51%	5.09%	2.30%
S	6.03%	3.61%	1.02%	-1.29%	7.83%	5.37%	2.64%	0.30%	5.73%	3.35%	0.70%	-1.57%	7.51%	5.09%	2.30%
W	-2.17%	-4.40%	0.62%	-1.68%	-0.18%	-2.46%	2.54%	0.20%	-2.34%	-4.54%	0.40%	-1.86%	-0.38%	-2.62%	2.30%
X	8.79%	6.31%	3.65%	1.28%	10.63%	8.11%	5.31%	2.91%	5.73%	3.35%	0.70%	-1.57%	7.51%	5.09%	2.30%
Y	-1.06%	-3.32%	1.02%	-1.29%	0.62%	-1.68%	2.64%	0.30%	-1.34%	-3.56%	0.70%	-1.57%	0.32%	-1.94%	2.30%

Note: Sim2222 refers to where all the parameters are perturbed by two standard errors.
Sim0022 refers to where γ and A is perturbed but α and β is at its estimated value.

Table 3.19 summarizes these results and compares them to the model's confidence intervals. Depending on the parameter set used in the sensitivity analysis, the range of results can be larger or smaller than the true confidence interval. Moreover, the sheer number of simulations that must be performed under sensitivity analysis impedes the analysis.

Table 3.19: Comparison between Confidence Intervals and Sensitivity Analysis Results

Variables	Confidence Interval	Sensitivity Analysis	
		Max	Min
C	4.68%	7.51%	0.30%
J	1.46%	0.40%	0.00%
L	8.96%	7.91%	0.00%
P_K	2.76%	9.20%	1.57%
Q	5.82%	9.20%	0.70%
S	4.68%	7.83%	0.30%
W	3.01%	2.54%	0.20%
X	7.05%	8.79%	0.70%
Y	2.49%	2.64%	0.32%

3.7 SELECTIVE COVARIANCE METHOD

The previous section shows that although sensitivity analysis is a cumbersome and ad hoc approach for determining the precision of a model's output, it does provide one worthwhile benefit that confidence intervals alone do not shows which of the parameter uncertainties are most important for the model's results. In other words, it provides feedback on where additional work

on parameterization is most needed. The procedure used for computing confidence intervals, however, can be adapted to provide similar information. The remainder of this section describes one possible approach, which will be referred to as the "selective covariance method".

Roughly speaking, the idea behind the selective covariance method is to perturb the variance-covariance matrix and then look at the effect of the perturbation on the model's confidence intervals. For example, if the value of the α parameter were known with perfect certainty, then all the elements of the first row and column of the parameter variance-covariance matrix would be zero. To test the importance of uncertainty about the true value of α, therefore, the confidence intervals could be calculated using a modified variance-covariance matrix in which the first row and column had been set to zero. If that caused the confidence intervals to be significantly different from those for the base case, it would suggest that uncertainty about α was relatively important.

To explore this issue, confidence intervals for twenty cases with different variance-covariance matrices were constructed.[35] Table 3.17 lists all of the variance-covariance matrices used and Table 3.18 shows the confidence intervals for each corresponding case. Case 9, for example, assumes that the estimates of γ and A are known with certainty. The resulting confidence intervals differ significantly from those for the base case. The confidence interval for labor supply becomes much narrower, changing from 8.9 percent to 1.9 percent. This result, combined with results from other cases (in particular, Case 11), points to

[35] Although the amount of data appears large, it should be noted that the actual calculations present no significant computational burden: all that is required is the multiplication of several small matrices.

an important conclusion: the estimates of γ and A are not precisely determined (having large standard errors) and uncertainty about those parameters is a major determinant of the model's overall precision. In addition, the covariance between the estimates of γ and A are nontrivial. Better estimation of these parameters would be very valuable.

Table 3.20: Selective Covariance Method

CASE 1:Full covariance

1.033E-05	-1.093E-05	2.207E-05	-5.163E-05
-1.093E-05	1.232E-05	-2.392E-05	5.760E-05
2.207E-05	-2.392E-05	5.990E-05	-1.161E-04
-5.163E-05	5.760E-05	-1.161E-04	2.711E-04

CASE 2: Certainty in ALPHA parameter

0	0	0	0
0	1.232E-05	-2.392E-05	5.760E-05
0	-2.392E-05	5.990E-05	-1.161E-04
0	5.760E-05	-1.161E-04	2.711E-04

CASE 3: Certainty in BETA parameter

1.033E-05	0	2.207E-05	-5.163E-05
0	0	0	0
2.207E-05	0	5.990E-05	-1.161E-04
-5.163E-05	0	-1.161E-04	2.711E-04

CASE 4: Certainty in ATECH parameter

1.033E-05	-1.093E-05	2.207E-05	0
-1.093E-05	1.232E-05	-2.392E-05	0
2.207E-05	-2.392E-05	5.990E-05	0
0	0	0	0

CASE 5: Certainty in ATECH parameter

1.033E-05	-1.093E-05	2.207E-05	0
-1.093E-05	1.232E-05	-2.392E-05	0
2.207E-05	-2.392E-05	5.990E-05	0
0	0	0	0

CASE 6: DIAGONAL covariance

1.033E-05	0	0	0
0	1.232E-05	0	0
0	0	5.990E-05	0
0	0	0	2.711E-04

CASE 7: DIAGONAL covariance with certainty in ALPHA

0	0	0	0
0	1.232E-05	0	0
0	0	5.990E-05	0
0	0	0	2.711E-04

CASE 8: DIAGONAL covariance with certainty in BETA

1.033E-05	0	0	0
0	0	0	0
0	0	5.990E-05	0
0	0	0	2.711E-04

CASE 9: DIAGONAL covariance with certainty in GAMMA

1.033E-05	0	0	0
0	1.232E-05	0	0
0	0	0	0
0	0	0	2.711E-04

CASE 10: DIAGONAL covariance with certainty in ATECH

1.033E-05	0	0	0
0	1.232E-05	0	0
0	0	5.990E-05	0
0	0	0	0

CASE 11: DIAGONAL covariance with certainty in GAMMA & ATECH

1.033E-05	0	0	0
0	1.232E-05	0	0
0	0	0	0
0	0	0	0

CASE 12: DIAGONAL covariance with certainty in BETA & ATECH

1.033E-05	0	0	0
0	0	0	0
0	0	5.990E-05	0
0	0	0	0

CASE 13: DIAGONAL covariance with certainty in BETA & GAMMA

1.033E-05	0	0	0
0	0	0	0
0	0	0	0
0	0	0	2.7109E-04

CASE 14: DIAGONAL covariance with certainty in BETA & GAMMA & ATECH

1.033E-05	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

CASE 15: DIAGONAL covariance with certainty in ALPHA & ATECH

0	0	0	0
0	1.232E-05	0	0
0	0	5.990E-05	0
0	0	0	0

CASE 16: DIAGONAL covariance with certainty in ALPHA & GAMMA

0	0	0	0
0	1.232E-05	0	0
0	0	0	0
0	0	0	2.7109E-04

CASE 17: DIAGONAL covariance with certainty in ALPHA & GAMMA & ATECH

0	0	0	0
0	1.232E-05	0	0
0	0	0	0
0	0	0	0

CASE 18: DIAGONAL covariance with certainty in ALPHA & BETA

0	0	0	0
0	0	0	0
0	0	5.990E-05	0
0	0	0	2.7109E-04

CASE 19: DIAGONAL covariance with certainty in ALPHA & BETA & ATECH

0	0	0	0
0	0	0	0
0	0	5.990E-05	0
0	0	0	0

CASE 19: DIAGONAL covariance with certainty in ALPHA & BETA & GAMMA

0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	2.7109E-04

Table 3.21: Confidence Intervals with Selective Covariance Method

CASE	c	j	l	pk	q	s	w	x	y
Residual	28.65%	0.59%	3.39%	3.28%	4.19%	28.65%	2.88%	7.02%	3.01%
Full Cov.	4.68%	1.46%	8.96%	2.76%	5.82%	4.68%	3.01%	7.05%	2.49%
2	4.44%	1.55%	9.49%	2.27%	4.44%	4.44%	3.17%	4.44%	2.74%
3	3.38%	1.20%	7.35%	4.42%	4.42%	3.38%	4.74%	5.60%	3.98%
4	6.82%	1.46%	8.96%	4.45%	8.00%	6.82%	1.33%	9.26%	0.95%
5	6.82%	1.46%	8.96%	4.45%	8.00%	6.82%	1.33%	9.26%	0.95%
6	5.84%	1.32%	8.08%	6.27%	6.03%	5.84%	4.04%	6.49%	3.47%
7	5.83%	1.32%	8.06%	6.08%	5.83%	5.83%	4.04%	5.83%	3.46%
8	5.62%	1.29%	7.88%	5.81%	5.81%	5.62%	3.57%	6.29%	3.08%
9	2.80%	0.31%	1.90%	3.62%	3.18%	2.80%	2.97%	3.98%	2.80%
10	5.37%	1.32%	8.08%	5.84%	5.58%	5.37%	3.33%	6.07%	2.61%
11	1.61%	0.31%	1.90%	2.81%	2.21%	1.61%	1.89%	3.26%	1.61%
12	5.13%	1.29%	7.88%	5.34%	5.34%	5.13%	2.74%	5.86%	2.06%
13	2.31%	0.09%	0.58%	2.76%	2.76%	2.31%	2.30%	3.65%	2.31%
14	0.28%	0.09%	0.58%	1.53%	1.53%	0.28%	0.18%	2.85%	0.28%
15	5.36%	1.32%	8.06%	5.63%	5.36%	5.36%	3.32%	5.36%	2.59%
16	2.79%	0.30%	1.81%	3.28%	2.79%	2.79%	2.97%	2.79%	2.79%
17	1.59%	0.30%	1.81%	2.35%	1.59%	1.59%	1.89%	1.59%	1.59%
18	5.61%	1.28%	7.86%	5.61%	5.61%	5.61%	3.57%	5.61%	3.07%
19	5.12%	1.28%	7.86%	5.12%	5.12%	5.12%	2.74%	5.12%	2.05%
20	2.29%	0.00%	0.00%	2.29%	2.29%	2.29%	2.29%	2.29%	2.29%

3.8 CHAPTER SUMMARY

In this chapter, confidence intervals of a small static computable general equilibrium model were derived and calculated. The procedure for calculating the confidence interval involves linearizing the model's equations and then using information about the variance-covariance matrices for the parameter estimates and residuals to calculate the variance of the model's endogenous variables. The procedure is straightforward and scales linearly with the number of parameters in the model, unlike Monte-Carlo simulation or other approaches.

In addition, confidence intervals that account for only uncertainty in the estimated parameters were compared to the overall confidence intervals. For some endogenous variables, uncertainty about parameters was the most important determinant of the confidence interval; for others, it was the unexplained variation in the estimating equations.

The chapter also examined how the confidence intervals calculated by the proposed method compare to the results obtained via Monte-Carlo simulation and various types of sensitivity analysis. The confidence intervals were found to be very similar to Monte-Carlo results despite being much easier to compute. They also dominate sensitivity analysis: they are easier to compute and interpret, and they can incorporate cross correlations in the parameter estimates and residuals.

Moreover, the chapter also explained how the GEMPACK software package can be used to calculate a nonlinear base case solution for a linearized model and proposed a very tractable method for determining which uncertainties in the parameter estimates have the largest effect on the confidence intervals.

The model used in this chapter, PROTO, was deliberately kept very simple to allow the analysis and exposition to be as clear as possible. The next chapter presents a much more realistic model of the US economy. It also extends the methodology to calculate confidence intervals for differences between simulations.

Chapter 4: An Intertemporal General Equilibrium Model

4.1 OVERVIEW OF THE INTERTEMPORAL MODEL STRUCTURE

This chapter presents a small econometric intertemporal general equilibrium model of the US economy that can be used to examine a realistic policy question. The production specification is discussed in section 4.1.1. Section 4.1.2 presents the household specification and section 4.1.3 presents the government specification. The rest of the equations and the market clearing conditions are presented in section 4.1.4. Section 4.2 derives an appropriate intertemporal welfare measure for the model. The rest of the chapter focuses on solving the steady state and intertemporal versions of the model, and on generating the confidence intervals.

4.1.1 Producer Model

The economy is basically divided into three aggregate industries: Energy, Materials, and Capital goods. The aggregation scheme for the industries and the definition of each industry is presented in Appendix A.

Each industry produces a distinct output (Q_i). Industries are modeled as competitive firms; which minimize total costs by choosing optimal quantities of primary inputs (labor, capital services) and intermediate goods (energy, and materials). Each firm is subject to a technology constraint represented by a CES production function, shown in equation 4.1.

$$Q_i(X_{it},\gamma,\sigma,A)=A_i\left(\gamma_{iK}^{\frac{1}{\sigma_i}}X_{iKt}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iL}^{\frac{1}{\sigma_i}}X_{iLt}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iE}^{\frac{1}{\sigma_i}}X_{iEt}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iM}^{\frac{1}{\sigma_i}}X_{iMt}^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}} \qquad (4.1)$$

Variable X_{ijt} is the demand by industry i for input j at time t. For example, X_{iKt}, and X_{iLt} are the capital services and labor demanded by industry i. Parameter γ_{ij} is a weighting parameter specific to industry i and input j. The parameter A_i reflects the level of technology in industry i and is constant over time. Finally, σ_i is elasticity of substitution for industry i.

Each industry faces the same input prices: p_K, p_L, p_E and p_M, the rental the price of capital, wage rate, producer price of energy, and producer price of materials, respectively. Sales taxes may be imposed on inputs of energy and materials, and are represented by τ_E and τ_M. Each firm's optimization at a given time can be written as:

$$min\ \ p_K X_{iK}+p_L X_{iL}+p_E(1+\tau_E)X_{iE}+p_M(1+\tau_M)X_{iM}$$

subject to

$$Q_i=A_i\left(\gamma_{iK}^{\frac{1}{\sigma_i}}X_{iK}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iL}^{\frac{1}{\sigma_i}}X_{iL}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iE}^{\frac{1}{\sigma_i}}X_{iE}^{\frac{\sigma_i-1}{\sigma_i}}+\gamma_{iM}^{\frac{1}{\sigma_i}}X_{iM}^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}} \qquad (4.2)$$

Purchaser's prices (after taxes) will henceforth be denoted by a circumflex. For example, the purchaser's price for energy is denoted as $\hat{p}_E$ and equals the producer price of energy, p_E, multiplied by the term $(1+\tau_E)$. For clarity of

notation, the time subscript will be dropped from now on unless otherwise mentioned. The Lagrangian is shown in equation 4.3:

$$L_i = p_K X_{iK} + p_K X_{iK} + \hat{p}_E X_{iE} + \hat{p}_M X_{iM} + \lambda\left\{Q_i - A_i\left(\sum_{j\in\{K,L,E,M\}} \gamma_{iji}^{\frac{1}{\sigma_i}} X_{ij}^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}}\right\} \tag{4.3}$$

Taking first-order conditions and rearranging produces the following factor demand equations:

$$X_{iK} = \gamma_{iK} p_K^{-\sigma_i} \frac{Q_i}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{\sigma_i}{1-\sigma_i}} \tag{4.4}$$

$$X_{iL} = \gamma_{iL} p_L^{-\sigma_i} \frac{Q_i}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{\sigma_i}{1-\sigma_i}} \tag{4.5}$$

$$X_{iE} = \gamma_{EL} p_E^{-\sigma_i} \frac{Q_i}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{\sigma_i}{1-\sigma_i}} \tag{4.6}$$

$$X_{iM} = \gamma_{iM} p_M^{-\sigma_i} \frac{Q_i}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{\sigma_i}{1-\sigma_i}} \tag{4.7}$$

Substituting the factor demand equations 4.4 through 4.7 into the accounting expression for total cost and simplifying produces the unit cost function for industry *i*:

$$c_i = \frac{1}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{1}{1-\sigma_i}} \tag{4.8}$$

Firms are assumed to earn zero profits. Under this restriction, the total value of output will equal the total cost of the inputs. This can be formulated as:

$$Q_i c_i = p_i Q_i \tag{4.9}$$

where p_i is the unit output price and c_i represents the unit cost. Inserting the cost function, equation 4.8, allows the zero profit condition for industry i to be expressed as a price frontier:

$$p_i = \frac{1}{A_i}\left(\gamma_{iK} p_K^{1-\sigma_i} + \gamma_{iL} p_L^{1-\sigma_i} + \gamma_{iE} \hat{p}_E^{1-\sigma_i} + \gamma_{iM} \hat{p}_M^{1-\sigma_i}\right)^{\frac{1}{1-\sigma_i}} \tag{4.10}$$

Using equation 4.10, the demand equations 4.4 through 4.7 can be simplified further. Demand for input j by industry i under zero profits can be expressed as:

$$X_{ij} = \gamma_{ij} A_i^{\sigma-1}\left(\frac{p_i}{p_j}\right)^{\sigma} Q_i \qquad i = \{E, M, I\}\ ,\ j = \{K, L, E, M\} \tag{4.11}$$

Thus, the optimal operation of an industry is expressed with five equations: a zero profit condition, and four demand equations. In total, fifteen equations define the production side of the model.

4.1.2 Consumer Model

Household behavior is modeled by a single infinitely-lived, forward-looking representative agent having perfect-foresight. The household consumes energy and material goods, and also demands labor and capital services. It receives income from supplying labor and capital services. In addition, it receives a lump-sum subsidy transfer from the government.

The household's labor supply is perfectly inelastic. Saving, on the other hand, is determined by intertemporal optimization and depends on present and future prices and interest rates.

Consumption is modeled as the outcome of a two-stage budgeting decision. The first stage is the *inter*period decision, in which the household allocates wealth (the present value of current and future income) across periods. In addition to determining each period's level of consumption, the interperiod model also determines the rate of saving. The decision is represented by an intertemporal utility function. The second stage is the *intra*period decision, where spending on particular goods in a given year is determined by an intratemporal utility function. It is represented by a felicity (or instantaneous utility) function.. The complete model, then, can be broken into two components which are linked by the model's intertemporal variables. The first component is referred to as *inter*period model and the second as the *intra*period model.

In the first stage of the optimization, the household maximizes an additively-separable intertemporal utility function of the form shown below

subject to its lifetime budget constraint.[36] The optimization problem can be written as:

$$\underset{C^a}{MAX} \qquad U_t = \int_t^{\infty} \ln\left(C^a(s)\right) e^{-\rho(s-t)} ds$$

subject to

$$W(t) = \int_t^{\infty} E(C^a(s), p(s)) e^{-\int_t^s r(v)dv} ds \tag{4.12}$$

where : ρ is the rate of time preference,

$C^a(s)$ is an index of consumption (discussed below) within period s,

$W(t)$ is wealth: the present value of current and future income,

$p(s)$ is a vector of prices at time s,

$E(C^a(s),p(s))$ is the expenditure required to purchase C^a, and

$r(v)$ is the instantaneous rate of interest at time v.

The Lagrangian for the problem can now be written:

$$L = \int_t^{\infty} \left(\ln C^a\right) e^{-\rho(s-t)} ds + \Lambda \left(W(t) - \int_t^{\infty} E(C^a, p) e^{-\int_t^s r(v)dv} ds \right) \tag{4.13}$$

Differentiating with respect to consumption gives:

[36] Additive separability of the intertemporal utility function enables us to formulate the intertemporal model as a two stage budgeting problem.

$$\frac{\partial L}{\partial C^a(s)} = \frac{1}{C^a(s)} e^{-\rho(s-t)} - \Lambda \left(\frac{\partial E(C^a, p)}{\partial C^a(s)} e^{-\int_t^s r(v)dv} \right) = 0 \tag{4.14}$$

Solving for consumption expenditure in year s gives:

$$C^a(s) \frac{\partial E(C^a, p)}{\partial C^a(s)} = \frac{1}{\Lambda} e^{\int_t^s (\rho - r(v))dv} \tag{4.15}$$

Within each period, utility is given by an index, C^a, of the household's consumption of goods and services. $E(C^a,p)$ is the corresponding expenditure function and is derived below. The consumption index, C^a, is represented by a Stone-Geary function:

$$C^a(s) = \prod_{i \in \{K,L,EM\}} (c_i(s) - \mu_i)^{\alpha_i} \tag{4.16}$$

where: $c_i(s)$ is consumption of good i at time s, and α_i and μ_i are parameters. Parameter μ_i is sometimes called the "subsistence level" of good i but formally it is an unrestricted parameter: it is not necessarily positive, as "subsistence level" would seem to imply.

The indirect utility function for 4.16 can be written as:

$$C^a = \left(E - \sum p_i(s)\mu_i \right) \prod_i \left(\frac{\alpha_i}{p_i} \right)^{\alpha_i} \tag{4.17}$$

Inverting equation 4.17 to obtain expenditure as a function of C^a and prices produces to the expression:

$$\begin{aligned} E = & \sum_{i \in \{K,L\}} p_i(s)\mu_i + \sum_{j \in \{E,M\}} \hat{p}_j(s)\mu_j \\ & + C^a(s) \prod_{i \in \{K,L\}} \left(\frac{p_i(s)}{\alpha_i} \right)^{\alpha_i} \prod_{j \in \{E,M\}} \left(\frac{\hat{p}_j(s)}{\alpha_j} \right)^{\alpha_j} \end{aligned} \tag{4.18}$$

Substituting this into 4.12, the present value of the lifetime budget constraint can be written as:

$$W(t) = \int_t^{\infty} \left\{ \begin{array}{l} \sum_{i \in \{K,L\}} p_i(s)\mu_i + \sum_{j \in \{E,M\}} \hat{p}_j(s)\mu_j + \\ C^a(s) \prod_{i \in \{K,L\}} \left(\frac{p_i(s)}{\alpha_i} \right)^{\alpha_i} \prod_{j \in \{E,M\}} \left(\frac{\hat{p}_j(s)}{\alpha_j} \right)^{\alpha_j} \end{array} \right\} e^{-\int_t^s r(v)dv} \, ds \tag{4.19}$$

For clarity of notation, a price index $\tilde{P}$ can be defined for C^a, allowing the expression to be written as:

$$W(t) = \int_t^{\infty} \left(\sum_{i \in \{K,L\}} p_i(s)\mu_i + \sum_{j \in \{E,M\}} \hat{p}_j(s)\mu_j + C^a(s)\tilde{P}(s) \right) e^{-\int_t^s r(v)dv} ds \tag{4.20}$$

where: *r(v)* is the interest rate and $\tilde{P}(s)$ is the price index for C^a, computed as the product of purchase prices weighted by the share parameter:

$$\tilde{P}(s) = \prod_{i \in \{K,L\}} \left(\frac{p_i(s)}{\alpha_i} \right)^{\alpha_i} \prod_{j \in \{E,M\}} \left(\frac{\hat{p}_j(s)}{\alpha_j} \right)^{\alpha_j} \tag{4.21}$$

Combined with equation 4.15, this equation allows the relationship between expenditure in two periods, *s* and *t*, to expressed as follows:

$$C^a(s)\tilde{P}(s) = C^a(t)\tilde{P}(t)e^{\int_t^s (\rho - r(v))dv} \tag{4.22}$$

Equations 4.18 and 4.22 allow the budget constraint to be rewritten as:

$$C^a(t)\tilde{P}(t) = \rho \left(W(t) - \int_t^{\infty} \left(\sum_{i \in \{K,L\}} p_i(s)\mu_i + \sum_{j \in \{E,M\}} \hat{p}_j(s)\mu_j \right) e^{-\int_t^s r(v)dv} ds \right) \tag{4.23}$$

The term on the left-hand side of equation 4.23 represents supernumerary expenditure at time *t* (expenditure above and beyond the subsistence levels of each good). As shown in equation 4.18, however, total expenditure at time *t* is the

sum of supernumerary expenditure and expenditure on subsistence levels. Using the price index for C^a, it can be written as:

$$E_t = C^a(t)\,\tilde{P}(t) + \sum_{i\in\{K,L\}} p_i(s)\mu_i + \sum_{j\in\{E,M\}} \hat{p}_j(s)\mu_j \tag{4.24}$$

For convenience, define a new function $\zeta(p,\mu)$ as:

$$\zeta(p,\mu) = \int_t^{\infty} \left(\sum_{i\in\{K,L\}} p_i(s)\mu_i(s) + \sum_{j\in\{E,M\}} \hat{p}_j(s)\mu_j(s) \right) e^{-\int_t^s r(v)dv}\, ds \tag{4.25}$$

Substituting 4.25 into 4.23 allows supernumerary expenditure to be written more compactly as:

$$C^a(t)\,\tilde{P}(t) = \rho\big(W(t) - \zeta(p,\mu)\big) \tag{4.26}$$

Thus, the present value of wealth at time *t* will be equal to the present value of the total expenditure, which is the sum of the committed expenditure on subsistence quantities plus the expenditure on supernumerary consumption. This stream of total expenditure at time *s* is represented as $E^s(s)$.[37]

[37] For computational purposes the expenditure and present value are made discrete and developed in the following way:
Error! Objects cannot be created from editing field codes.

Error! Objects cannot be created from editing field codes.

Error! Objects cannot be created from editing field codes.

In the second stage, the household decides on the optimal allocation of consumption across commodities given each period's expenditure by maximizing the instantaneous utility function given in equation 4.16 subject to the level of expenditure determined in the first stage.[38] The household intraperiod optimization problem can be stated as:

$$\max_{\{c_i\}} \quad C^a = \prod_{i \in \{K,L,EM\}} (c_i - \mu_i)^{\alpha_i}$$

subject to

$$p_K c_K + p_L c_L + \hat{p}_E c_E + \hat{p}_M c_M = E \tag{4.27}$$

The first order conditions for this problem result in the well-known linear expenditure system (LES) of demands as follows:

$$c_i = \mu_i + \frac{\alpha_i}{p_i}\left\{E_t - \sum_{i \in \{K,L\}} p_i \mu_i - \sum_{j \in \{E,M\}} \hat{p}_j \mu_j\right\} \quad , \quad i = \{K, L, E, M\} \tag{4.28}$$

or, expressed more compactly:

$$c_i = \mu_i + \frac{\alpha_i}{p_i}\left\{C^a(s)\,\tilde{P}(s)\right\} \quad , \quad i = \{K, L, E, M\}$$

[38] Alternatively, it could be described as minimizing the cost of achieving the value of C^a determined in the first stage.

where: $\left\{ \sum_{i \in \{K,L\}} p_i \mu_i + \sum_{j \in \{E,M\}} \hat{p}_j \mu_j \right\}$ is total subsistence expenditure

$\left\{ E_t - \sum_{i \in \{K,L\}} p_i \mu_i - \sum_{j \in \{E,M\}} \hat{p}_j \mu_j \right\}$ is "supernumerary" income

For the LES system, supernumerary income is allocated among all the commodities in fixed proportions given by the α_i parameters.

Household net income, *Y*, which includes the government transfer, *TR*, in the form of a lump sum subsidy, is:

$$Y = p_L (1 - \tau_L) L + p_K (1 - \tau_K) K + TR \tag{4.29}$$

Finally, the household owns the entire capital stock, which evolves according to the accumulation equation:[39]

$$\dot{K} = I - \delta K \tag{4.30}$$

where δ is the rate of depreciation.

The interest rate is determined by the following arbitrage equation that links the rental price of capital with the price of new investment:[40]

---

[39] The household owns the entire capital stock (*K*), which is demanded by the firms as well as the household. The amount consumed by the household is c_k less than the total capital stock ($C_K < K$).

[40] The investment goods and the capital stock are measured in different units in the US data that is constructed. To make the units consistent the capital stock is scaled by an adjusted parameter while solving the model. However, for ease of notation this discussion is omitted from the text. For a detailed discussion on the adjustment factor, see Wilcoxen (1988).

$$p_K(1-\tau_K)S^{PK} = (r+\delta)p_I \tag{4.31}$$

where S^{PK} is a constant implicitly determined by the normalizations used for the rental price of capital and the purchase price of new capital goods when the model's data set was constructed.

Saving is equal to income (Y) net of expenditures (E) at each time, and it is exclusively used to purchase new investment goods. Thus, the remainder of the equations for the consumer model, namely expenditure (E), savings (S), and investment (I), are given as:

$$E = p_K c_K + p_L c_L + \hat{p}_E c_E + \hat{p}_M c_M \tag{4.32}$$

$$S = Y - E \tag{4.33}$$

$$S = p_I I \tag{4.34}$$

The intertemporal model consists of two differential equations: equation 4.30, which describes the evolution of the capital stock, and the differential form of 4.19, which describes the evolution of wealth.[41] Two constants of integration are required to tie down the solution path. Hence, two boundary conditions are needed. One boundary condition that arises naturally is the initial capital stock, K_0, which is observable at the initial period.[42] The other boundary condition is

[41] An equivalent approach would be to use the model's Euler equation as the second differential equation.

[42] Unlike the shadow prices and co-state variables that are unobservable in the initial period.

imposed by assumption: that as time becomes large, the economy will approach its steady state.[43] This assumption ensures that in such a state the rate of change of all the variables to be zero. To impose this condition, equation 4.30 is set to zero at the steady state. Problems boundary conditions split between an initial and a terminal time are known as the "two point boundary value" problem. Thus, the boundary conditions are written as:

$$K(0) = K_0 \tag{4.35}$$

$$I^{SS} = \delta K^{SS} \tag{4.36}$$

where K_0, I^{SS}, and K^{SS} are initial capital stock, steady state investment and the steady state capital stock. This completes the consumer model.

4.1.3 Government

In order to keep the model as transparent as possible, the government collects taxes and returns the revenue as lump sum rebate to the household. The government imposes the following taxes:

(i) Sales tax on energy and materials

(ii) Proportional tax on labor and capital income

[43] In particular, it is assumed that as agents form expectations about events further and further into the future, they do not expect that the economy will move increasingly far from its steady state.

The government runs a balanced budget at all times and total tax revenue is transferred to the consumer via the lump sum subsidy. The government budget constraint is written as:

$$TR = \tau_L p_L L + \tau_K p_K K + \tau_E p_E \left(c_E + X_{ME} + X_{IE}\right) + \tau_M p_M \left(c_M + X_{EM} + X_{IM}\right) \tag{4.37}$$

4.1.4 The Rest of the Equations and Equilibrium Conditions

The rest of the equations for the model are a set of market clearing conditions and accounting equations. For energy and materials, total supply must equal the sum of all intermediate demands plus the demand by households. Also, investment must equal the supply of new capital goods. In addition, there are market clearing conditions for the primary input factors: capital and labor. As a group, the market clearing equations are as follows:

$$X_{EE} + X_{ME} + c_E + X_{IE} = Q_E \tag{4.38}$$

$$X_{EM} + X_{MM} + c_M + X_{IM} = Q_M \tag{4.39}$$

$$I = Q_I \tag{4.40}$$

$$X_{EK} + X_{MK} + c_K + X_{IK} = K \tag{4.41}$$

$$X_{EL} + X_{ML} + c_L + X_{IL} = L \tag{4.42}$$

The general equilibrium solution of the model is a set of prices for all goods and services at all points in time such that all budget constraints are satisfied and all markets clear in all periods.

4.2 THE STEADY STATE

The steady state for this model is straightforward to construct and computationally easy to solve. To convert the full model to a steady state version, the rates of change of the state variable (the capital stock) and the control variable (consumption) are set to zero. The dynamic equations of the model then simplify to the following expressions:

$$I = \delta K \tag{4.43}$$

$$W = \frac{E^s}{r} \tag{4.44}$$

All the other equations of the model remain the same in the steady state version. The steady state model is useful because it provides a benchmark for comparison and can be used to analyze the long run effect of the policy.

4.3 DATA AND ESTIMATION

The producer model consists of the three zero profit conditions and twelve demand equations given in equation 4.10 and 4.11. The demand equations require information on the output quantities along with the input and output prices. Since the demand equation is homogenous of degree one in the level of output, the output level can be eliminated from the demand equations by expressing them as share equations.

Multiplying equation 4.11 by the corresponding ratio of the input price to output price, p_j / p_i, allows the equation to be written as:

$$\omega_{ji} = \frac{X_{ij} p_j}{p_i Q_i} = \gamma_{ij} A_i^{\sigma-1} \left(\frac{p_i}{p_j} \right)^{\sigma} \quad \text{i} = \{\text{E, M, I}\}, \text{j} = \{\text{K, L, E, M}\} \qquad (4.45)$$

Rearranging this gives the expenditure share on commodity *i* for industry *j* in terms of exogenous variables (prices). For example, multiplying both sides of the materials demand equation for energy industry by p_M / p_E yields the share equation (ω_{ME}) in the form:

$$\omega_{ME} = \frac{p_M X_{ME}}{p_E Q_E} = \gamma_{EM} A_E^{\sigma_E - 1} \left(\frac{p_E}{p_M} \right)^{\sigma_E - 1} \qquad (4.46)$$

For the sake of clarity and continuity, the econometric equations used for estimation are presented in equation 4.47 through 4.53.[44]

$$p_E = \frac{1}{A_E} \left(\gamma_{EK} p_K^{1-\sigma_E} + \gamma_{EL} p_L^{1-\sigma_E} + \gamma_{EE} p_E^{1-\sigma_E} + \gamma_{EM} \hat{p}_M^{1-\sigma_E} \right)^{\frac{1}{1-\sigma_E}} + \varepsilon_E \qquad (4.47)$$

$$p_M = \frac{1}{A_M} \left(\gamma_{MK} p_K^{1-\sigma_M} + \gamma_{ML} p_L^{1-\sigma_M} + \gamma_{ME} \hat{p}_E^{1-\sigma_M} + \gamma_{MM} p_M^{1-\sigma_M} \right)^{\frac{1}{1-\sigma_M}} + \varepsilon_M \qquad (4.48)$$

$$p_I = \frac{1}{A_I} \left(\gamma_{IK} p_K^{1-\sigma_I} + \gamma_{IL} p_L^{1-\sigma_I} + \gamma_{IE} \hat{p}_E^{1-\sigma_I} + \gamma_{IM} \hat{p}_M^{1-\sigma_I} \right)^{\frac{1}{1-\sigma_I}} + \varepsilon_I \qquad (4.49)$$

[44] Note: For the energy and the materials sector it is assumed that the sectors do not pay sales taxes on their own products used as intermediate input in the production.

$$\omega_{iK} = \gamma_{iK}\left(A_i \frac{p_i}{p_K} \right)^{\sigma_i - 1} + \varepsilon_{iK} \quad , \quad i \in \{E, M, I\} \tag{4.50}$$

$$\omega_{iL} = \gamma_{iL}\left(A_i \frac{p_i}{p_L} \right)^{\sigma_i - 1} + \varepsilon_{iL} \quad , \quad i \in \{E, M, I\} \tag{4.51}$$

$$\omega_{iE} = \gamma_{iE}\left(A_i \frac{p_i}{\hat{p}_E} \right)^{\sigma_i - 1} + \varepsilon_{iE} \quad , \quad i \in \{E, M, I\} \tag{4.52}$$

$$\omega_{iM} = \gamma_{iM}\left(A_i \frac{p_i}{\hat{p}_M} \right)^{\sigma_i - 1} + \varepsilon_{iM} \quad , \quad i \in \{E, M, I\} \tag{4.53}$$

There are six parameters per industry for a total of eighteen parameters overall: twelve γ's, three σ's and three A's. Without loss of generality, the within-industry γ parameters are assumed to add up to unity. For example, the γ parameter for materials can be expressed in terms of the other γ parameters as:

$$\gamma_{EM} = 1 - (\gamma_{EK} + \gamma_{EL} + \gamma_{EE}). \tag{4.54}$$

The data used for this model is aggregated from the US input-output tables. The data set construction and aggregation procedures are explained in detail in Appendix A. The producer model, consisting of fifteen equations, is estimated simultaneously by nonlinear least squares using TSP. The estimated parameter and residual variance-covariance matrices are retained for confidence interval calculations. The parameter estimates are presented in Table 4.1.

The elasticities of substitution for all three firms are clearly not equal to one, although for investment it is very close to one. This implies that the production technology is not Cobb-Douglas. The energy industry consumes significant amounts of all the inputs; the materials sector consumes very little energy; and the investment sector principally uses labor and materials. The substitution elasticity is particularly low for the materials sector and is close to one for investment.

One of the main objectives of econometric estimation is to determine the variance-covariance matrices for confidence interval calculations. These matrices are shown in Table 4.2 and Table 4.3. In addition, the corresponding correlation matrices are displayed graphically in Figure 4.1 and Figure 4.2. Examining the two figures provide qualitative information about the relationships between the estimated parameters and the relationships between the residuals. Off-diagonal dark spots indicate absolute correlation between 0.5 to 1. Both figures indicate that there are substantial cross-correlations, underscoring the need for computing confidence intervals rather than doing traditional sensitivity analysis (in which parameters are tested independently and residuals are ignored). These figures alone are very informative and illustrate the importance of covariance matrices.

Table 4.1: Parameter Estimates for the Producer Model

Estimates of the Parameter	Industries		
	Energy (E)	Materials (M)	Investment (I)
γ_K	0.269 (0.0019)	0.16 (0.0010)	0.041 (0.0005)
γ_L	0.159 (0.0019)	0.368 (0.0017)	0.182 (0.0021)
γ_E	0.416 (0.0040)	0.032 (0.0004)	0.013 (0.0008)
γ_M	0.155 (0.00098)	0.440 (0.00094)	0.765 (0.0023)
σ	0.850 (0.0041)	0.243 (0.0097)	0.932 (0.0034)
A	1.22 (0.015)	0.99 (0.011)	0.91 (0.010)

Standard errors are given in parenthesis

Table 4.2: Producer Model Estimated Parameter Covariance Matrix

Industry Parameters		Energy						Materials			
		γ_K	γ_L	γ_E	γ_M	σ	A	γ_K	γ_L	γ_E	γ_M
Energy	γ_K	3.8E-06									
	γ_L	1.6E-06	3.6E-06								
	γ_E	-6.4E-06	-6.5E-06	1.6E-05							
	γ_M	9.9E-07	1.3E-06	-3.2E-06	9.6E-07						
	σ	1.1E-06	-1.8E-06	4.1E-07	2.7E-07	1.7E-05					
	A	1.2E-06	1.9E-05	-2.6E-05	6.3E-06	-3.1E-05	2.3E-04				
Materials	γ_K	-5.6E-07	-1.5E-06	2.5E-06	-4.5E-07	1.3E-06	-8.5E-06	1.1E-06			
	γ_L	2.1E-06	2.7E-06	-5.8E-06	1.1E-06	-1.5E-06	9.4E-06	-1.4E-06	3.0E-06		
	γ_E	-5.8E-07	-5.6E-07	1.4E-06	-2.6E-07	-3.6E-08	-4.3E-07	1.6E-07	-5.1E-07	1.8E-07	
	γ_M	-9.5E-07	-6.1E-07	1.9E-06	-3.7E-07	2.4E-07	-4.0E-07	6.9E-08	-1.1E-06	1.8E-07	8.8E-07
	σ	-2.2E-06	-5.9E-06	1.1E-05	-2.6E-06	8.1E-06	-6.0E-05	6.1E-06	-7.2E-06	-5.4E-07	1.7E-06
	A	2.2E-07	-1.1E-06	1.3E-06	-4.5E-07	1.2E-06	-1.0E-05	7.7E-07	-8.1E-07	2.0E-08	1.2E-08
Investment	γ_K	-4.6E-08	1.1E-07	-9.1E-08	2.6E-08	1.3E-08	-9.8E-07	1.2E-07	2.1E-07	-4.8E-08	-2.9E-07
	γ_L	1.1E-06	1.0E-06	-2.2E-06	1.1E-07	-1.6E-06	6.0E-07	-3.4E-07	1.7E-06	-3.1E-07	-1.1E-06
	γ_E	4.9E-07	-4.6E-10	-6.2E-07	1.3E-07	-4.0E-07	1.3E-06	-3.8E-08	1.7E-07	9.6E-09	-1.4E-07
	γ_M	-1.6E-06	-1.1E-06	2.9E-06	-2.7E-07	2.0E-06	-9.0E-07	2.5E-07	-2.1E-06	3.5E-07	1.5E-06
	σ	-7.2E-07	-1.6E-06	2.8E-06	-4.6E-07	1.2E-06	-3.0E-05	4.1E-07	-5.1E-07	-2.4E-07	3.4E-07
	A	2.0E-07	-1.2E-05	1.5E-05	-3.7E-06	2.1E-05	-1.4E-04	7.3E-06	-8.2E-06	7.7E-08	8.1E-07

Contd.

Producer Model Estimated Parameter Covariance Matrix

Industry Parameters		Investment							
		σ	A	γ_K	γ_L	γ_E	γ_M	σ	A
Energy	γ_K								
	γ_L								
	γ_E								
	γ_M								
	σ								
	A								
Materials	γ_K								
	γ_L								
	γ_E								
	γ_M								
	σ	9.5E-05							
	A	4.6E-06	1.2E-06						
Investment	γ_K	4.2E-08	7.3E-08	2.6E-07					
	γ_L	-2.8E-06	-7.7E-08	6.0E-07	4.5E-06				
	γ_E	-8.6E-07	9.6E-09	-1.4E-07	-5.9E-07	7.2E-07			
	γ_M	3.6E-06	-5.7E-09	-7.2E-07	-4.5E-06	3.5E-09	5.3E-06		
	σ	9.9E-06	7.1E-07	4.5E-08	-2.1E-07	-9.8E-07	1.1E-06	1.1E-05	
	A	4.9E-05	9.4E-06	5.4E-07	-2.9E-06	-1.3E-06	3.6E-06	1.7E-05	1.1E-04

Table 4.3: Producer Model Estimated Residual Covariance Matrix

	Eq1	Eq2	Eq3	Eq4	Eq5	Eq6	Eq7	Eq8	Eq9	Eq10	Eq11	Eq12	Eq13	Eq14	Eq15
Eq1	1.8E-03														
Eq2	3.4E-03	2.0E-02													
Eq3	-3.6E-03	-1.5E-02	1.8E-02												
Eq4	4.3E-04	-9.1E-04	-9.6E-04	2.5E-03											
Eq5	1.0E-04	1.3E-03	-4.1E-04	-5.9E-04	3.8E-04										
Eq6	1.6E-03	8.3E-03	-6.1E-03	-5.6E-04	5.0E-04	4.2E-03									
Eq7	-3.2E-03	-8.5E-03	1.2E-02	-2.4E-03	2.6E-04	-3.5E-03	1.2E-02								
Eq8	-1.5E-03	-8.2E-03	6.4E-03	6.4E-04	-8.1E-04	-3.7E-03	3.8E-03	4.4E-03							
Eq9	7.1E-05	7.9E-04	-9.0E-04	-5.4E-04	2.6E-04	3.2E-04	-1.4E-04	-5.5E-04	5.3E-04						
Eq10	1.2E-03	4.3E-03	-5.3E-03	6.4E-04	-8.9E-05	2.3E-03	-5.4E-03	-2.5E-03	-7.6E-04	8.0E-03					
Eq11	2.7E-03	7.3E-03	-7.8E-03	3.9E-03	-8.2E-04	5.7E-03	2.4E-03	-1.0E-03	-4.5E-04	-1.5E-02	2.1E-01				
Eq12	3.5E-05	-4.9E-04	2.3E-04	-5.2E-06	-8.1E-04	7.3E-05	-1.3E-03	1.2E-03	-9.0E-04	2.7E-03	-9.4E-03	6.2E-03			
Eq13	-1.1E-03	-5.7E-03	6.7E-03	7.4E-05	-4.7E-04	-1.4E-03	3.8E-03	2.7E-03	-4.0E-04	-2.6E-03	3.8E-03	7.7E-04	5.1E-03		
Eq14	1.6E-04	7.1E-04	-2.8E-04	-1.6E-04	1.1E-04	3.5E-04	-4.7E-05	-3.8E-04	2.5E-05	1.7E-04	2.0E-04	-1.1E-04	-1.8E-04	6.2E-05	
Eq15	2.4E-03	1.1E-02	-1.3E-02	-1.3E-03	9.1E-04	4.8E-03	-7.4E-03	-5.7E-03	1.1E-03	5.1E-03	5.4E-04	-5.8E-04	-6.0E-03	4.0E-04	1.2E-02

Figure 4.1: Contour Diagram for Estimated Parameter Correlation Matrix (Producer)

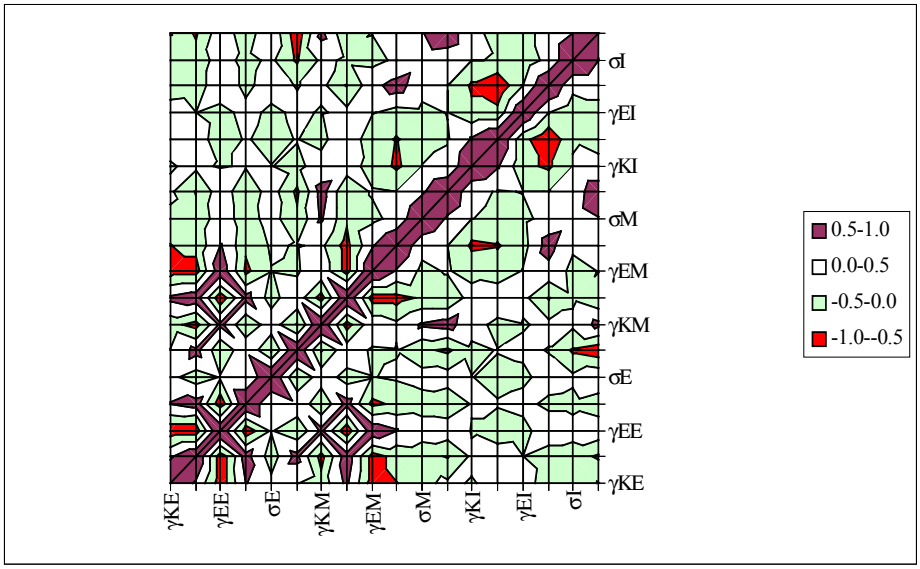
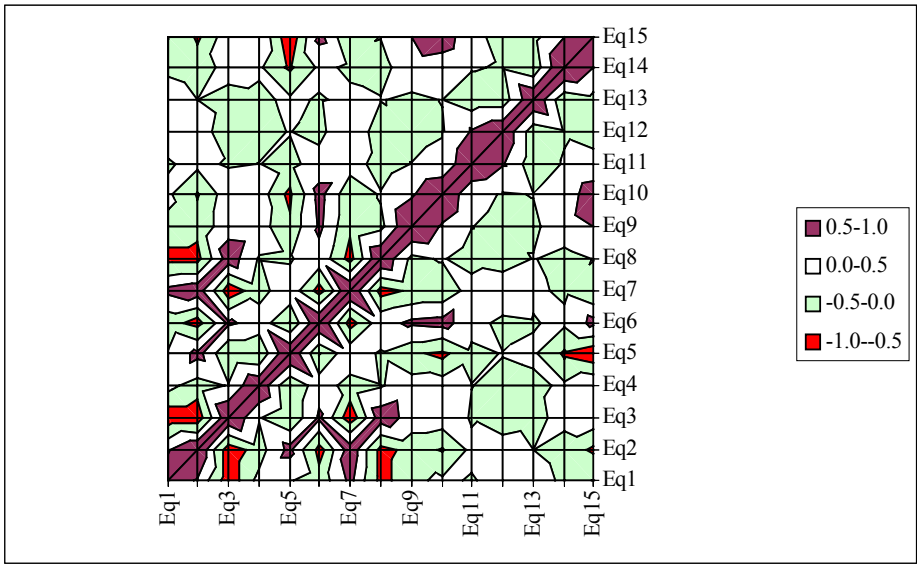


Figure 4.2: Contour Diagram for Estimated Residual Correlation Matrix (Producer)



The consumer model, on the other hand, consists of four equations expressed in equation 4.28 with eight parameters: four budget shares (α) that sum to unity and four subsistence quantity (μ) parameters. For the Stone-Geary utility function to be meaningful, the subsistence level must be restricted to be less than any actual level of consumption in all of the sample periods. It is formally stated in the form:

$$\left(c_{it} - \mu_{it}\right) \geq 0 \;\; \forall \;\; i \in \left\{K, L, E, M\right\} \; and \;\; \forall \;\; t \, .$$

To accommodate the above restriction each of the subsistence parameters are estimated indirectly using a logistic function with a new parameter (υ) that takes the form:

$$\mu_i = \frac{\min\left\{c_{it}\right\}}{1+\upsilon_i^2} \quad , \; i \in \left\{K, L, E, M\right\}. \tag{4.55}$$

If the estimated value of υ is very small then subsistence quantity is approximately equal to the minimum level of consumption. However, if υ takes on a large value, then the subsistence level is much smaller than the minimum consumption level.[45] This guarantees that the subsistence parameters are smaller that the minimum consumption level. The estimated values of the subsistence

[45] There is no restriction on the sign of the subsistence level, that is, it could theoretically be negative. However, the assumption of the logistic function restricts the subsistence parameters to be positive.

parameters can be back calculated given the estimated υ parameters.[46] Substituting equation 4.55 into equation 4.28 gives the model used for estimation. Again for the sake of completeness, the econometric equations for the consumer side are given as:

$$c_K = \frac{\min\{c_K\}}{1+\upsilon_K^2} + \frac{\alpha_K}{p_K}\left\{Y - p_I I - \sum_{i\in\{K,L\}} p_i \frac{\min\{c_i\}}{1+\upsilon_i^2} - \sum_{j\in\{E,M\}} \hat{p}_j \frac{\min\{c_j\}}{1+\upsilon_j^2}\right\} + \varepsilon_K \tag{4.56}$$

$$c_L = \frac{\min\{c_L\}}{1+\upsilon_{LK}^2} + \frac{\alpha_L}{p_L}\left\{Y - p_I I - \sum_{i\in\{K,L\}} p_i \frac{\min\{c_i\}}{1+\upsilon_i^2} - \sum_{j\in\{E,M\}} \hat{p}_j \frac{\min\{c_j\}}{1+\upsilon_j^2}\right\} + \varepsilon_L \tag{4.57}$$

$$c_E = \frac{\min\{c_E\}}{1+\upsilon_E^2} + \frac{\alpha_E}{p_E}\left\{Y - p_I I - \sum_{i\in\{K,L\}} p_i \frac{\min\{c_i\}}{1+\upsilon_i^2} - \sum_{j\in\{E,M\}} \hat{p}_j \frac{\min\{c_j\}}{1+\upsilon_j^2}\right\} + \varepsilon_E \tag{4.58}$$

$$c_M = \frac{\min\{c_M\}}{1+\upsilon_M^2} + \frac{\alpha_M}{p_M}\left\{Y - p_I I - \sum_{i\in\{K,L\}} p_i \frac{\min\{c_i\}}{1+\upsilon_i^2} - \sum_{j\in\{E,M\}} \hat{p}_j \frac{\min\{c_j\}}{1+\upsilon_j^2}\right\} + \varepsilon_M \tag{4.59}$$

The cross-equation restrictions for the system are implied by the integrability conditions. In order to avoid singularity in the covariance matrix one of the share equations from the system is dropped. The remaining equations identify all of the parameters and were estimated together by full information maximum likelihood (FIML). The estimates obtained are presented in Table 4.4.[47]

[46] TSP can calculate this by the function analyz.

[47] Either FIML or a variant of generalized least squares (GLS) methods can be used. See Mansur and Walley (1984) for detailed discussion.

Table 4.4: Transformed Consumer Model Parameter Estimates

Estimated Parameter	Estimate	Standard Error
υ_K	0.17E-07	12.6E+08
υ_L	0.369	0.5540
υ_E	-0.695E-06	0.10E+07
υ_M	0.419	0.2940
α_M	0.6314	0.0271
α_K	0.1990	0.0859
α_L	0.127	0.0123

Substituting the estimated υ parameters into equation 4.66, the estimates of the subsistence parameters, μ, and the parameter covariance matrix can be recovered. The estimates obtained are presented in Table 4.5.

The energy share is very small and not estimated precisely. Also, the subsistence level for energy is not estimated precisely. This imprecise estimation and a large standard error for energy parameter will be subsequently referred to in discussing the solution. Subsistence levels for energy and capital consumption are effectively constrained to their boundary points, that is, to the minimum levels in the dataset.

Table 4.5: Consumer Model Parameter Estimates

Estimated Parameter	Estimate	Standard Error	Minimum level of consumption (billion units)
μ_E	52,760	77,977	52,760
μ_M	674,032	141,749	792,539
μ_K	55,430	23,730	55,430
μ_L	107,950	38,859	122,655
α_E	0.0423	0.0355	-
α_M	0.6314	0.0271	-
α_K	0.1990	0.0859	-
α_L	0.127	0.0123	-

As in the producer model estimation, parameter and residual variance-covariance matrices for the consumer model are recovered for calculating confidence intervals. Again, the contours in Figure 4.3 show that there are strong correlations between many of the parameter estimates. In addition, there is a strong correlation between the residuals in the capital services and materials equations.

Table 4.6: Consumer Model Estimated Parameter Covariance Matrix

	μE	μM	μK	μL	α_L	α_E	α_M	α_M
μE	6.08E+09							
μM	-6.08E+09	2.01E+10						
μK	-4.68E+08	2.61E+09	5.63E+08					
μL	-2.24E+09	5.00E+09	6.63E+08	1.51E+09				
α_L	8.86E+02	-1.04E+03	-1.08E+02	-4.08E+02	1.51E-04			
α_E	-2.75E+03	3.13E+03	2.75E+02	1.09E+03	-4.05E-04	1.26E-03		
α_M	1.93E+03	-2.92E+03	-2.12E+02	-8.34E+02	2.69E-04	-8.99E-04	7.37E-04	
α_M	-6.46E+01	8.27E+02	4.47E+01	1.56E+02	-1.51E-05	4.76E-05	-1.06E-04	7.38E-05

Figure 4.3: Contour Diagram for Estimated Parameter Correlation Matrix (Consumer)

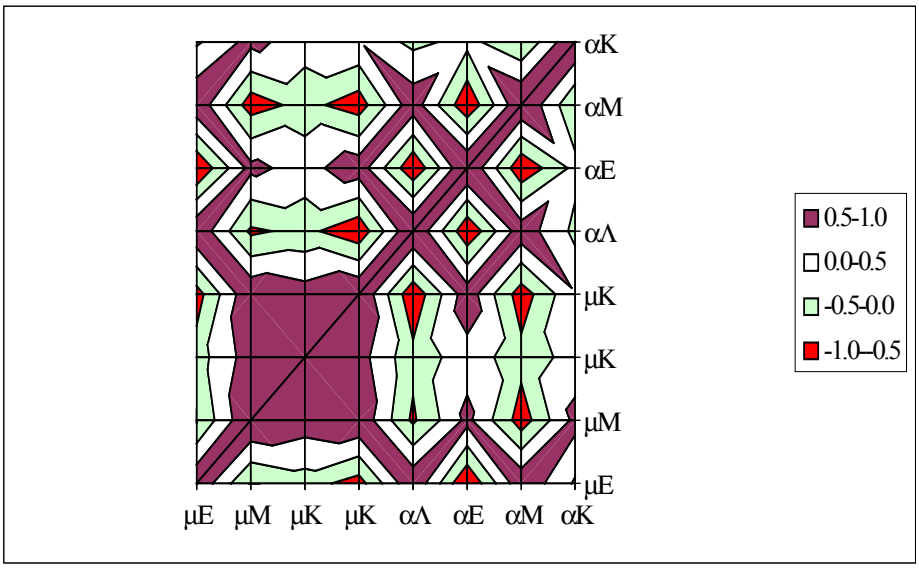
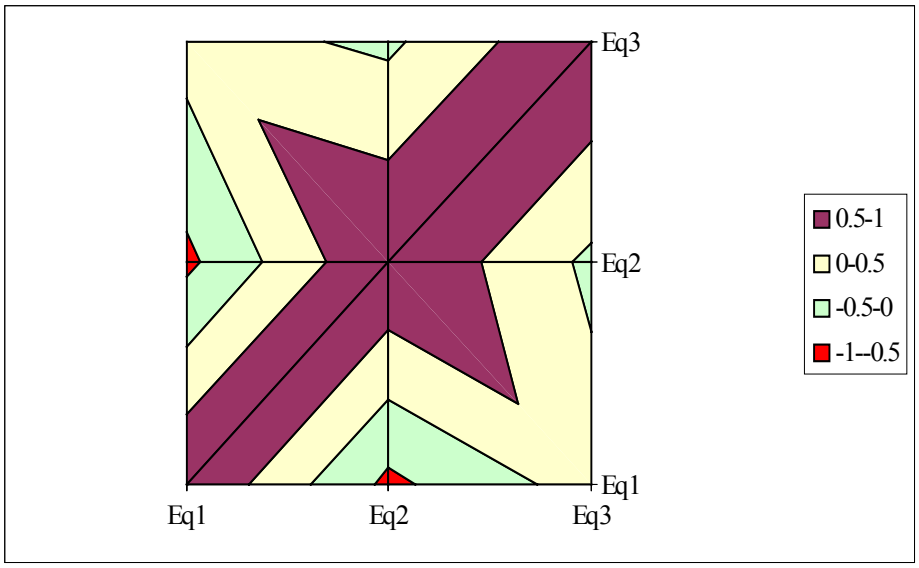


Table 4.7: Consumer Model Estimated Residual Covariance Matrix

	Eq1	Eq2	Eq3
Eq1	1.26E+08		
Eq2	-4.81E+07	4.84E+07	
Eq3	1.45E+07	-4.06E+06	3.51E+07

Figure 4.4: Contour Diagram for Consumer Residual Correlation Matrix



4.4 RESULTS

4.4.1 Steady State Results

The solution to the steady state model is presented in this section. The model is solved assuming that the economy converges to the steady state by the year 2040, that is, approximately sixty years from the benchmark year of 1982.

The model is thus simulated for sixty years into the future. The exogenous variables for the steady state are the tax rates and labor supply. The wage rate is also exogenous because it is the numeraire for the model.

The results for key variables are given in Table 4.8.[48] In addition to being of interest in themselves, these values will later be used in building the base case for the full intertemporal model.

Table 4.8: Base Case Solution

Variable	Description	Base case	Units
PI	Price of investment goods	$1.1029	dollars
PK	Rental price of capital	$0.7345	dollars
PE	Price of energy good	$0.6453	dollars
PM	Price of materials	$1.0009	dollars
Y	Full income	$3,110,284	mil.
R	Interest rate	$0.0288	dollars
QK	Capital stock	13,500,520	mil.
QE	Quantity of energy	641,662	mil.
QM	Quantity of materials	4,795,288	mil.
QI	Quantity of Investment	621,024	mil.
S	Subsidy	$839,909	mil.
W	Present value of wealth	$77,728,840	mil.
E	Expenditure on consumption	$3,110,284	mil.
Sa	Savings	$684,927	mil.

4.4.1.1 Confidence Intervals with Parameter Uncertainty

Confidence intervals for the steady state are calculated by the approach described in Chapter 3 Section 3.3.1, taking parameter uncertainty into account. The confidence intervals are presented in Table 4.9 as both percentage deviations

[48] The results of only few variables, that are deemed important, will be presented. The steady state alone has forty eight variables.

from the base case, and as absolute levels. Figure 4.5 gives a graphical comparison of the confidence intervals for the key endogenous variables.

Table 4.9: Base Case Confidence Intervals*

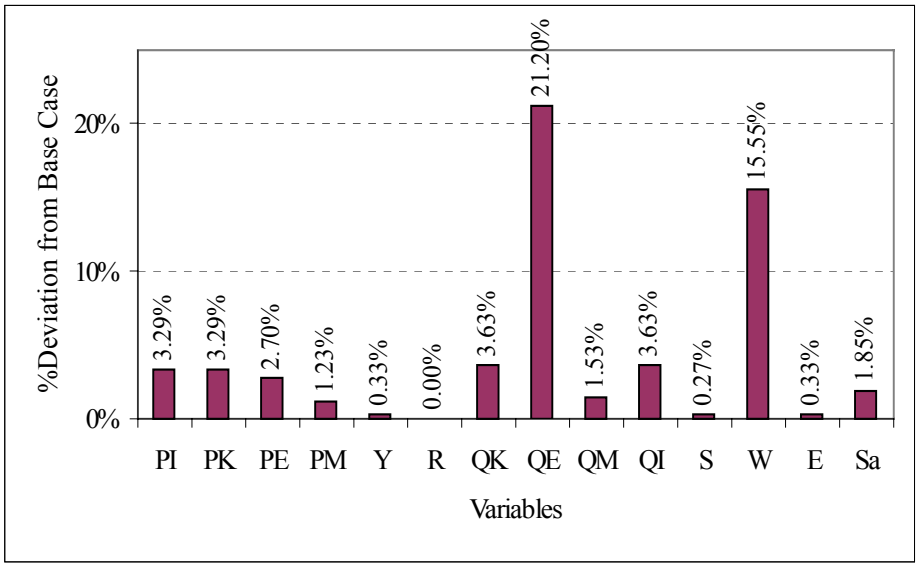
Variables	Lower Bound	Base Case	Upper Bound	Confidence Intervals
PI	$1.0666	$1.1029	$1.1392	3.29%
PK	$0.7103	$0.7345	$0.7586	3.29%
PE	$0.6279	$0.6453	$0.6627	2.70%
PM	$0.9886	$1.0009	$1.0131	1.23%
Y	$3,100,020	$3,110,284	$3,120,548	0.33%
R	$0.0288	$0.0288	$0.0288	0.00%
QK	13,009,800	13,500,520	13,991,240	3.63%
QE	505,622	641,662	777,702	21.20%
QM	4,722,145	4,795,288	4,868,431	1.53%
QI	598,481	621,024	643,567	3.63%
S	$842,177	$839,909	$842,277	0.27%
W	$65,642,005	$77,728,840	$89,815,675	15.55%
E	$3,100,020	$3,110,284	$3,120,548	0.33%
Sa	$672,256	$684,927	$697,598	1.85%

*See Table 4.8 for units

Accounting for parameter uncertainty, the quantity of energy and total wealth are the least precise of all the variables, having confidence intervals of 21 and 16 percent, respectively. The large uncertainty in energy is due to the imprecise estimates of the energy parameters. On the other hand, consumer expenditure, the lump-sum subsidy and income are all determined with surprising precision: the confidence intervals are less than half a percent. The confidence intervals for investment and the capital stock (shown in the table as QK) deviate by the same amount, a direct result of the steady-state condition. A key result is

the zero-width confidence interval for the interest rate. The reason is that in the steady state, the interest rate must exactly equal the time preference rate and the time preference rate itself was imposed rather than being estimated. Had the time preference rate been estimated, its confidence interval would have carried through to the interest rate.

Figure 4.5: Confidence Intervals in Percent Deviation



4.4.1.2 Confidence Intervals with Residual Uncertainty

As discussed in previous chapters, the confidence intervals for projected variables (which include the effect of estimation residuals) are generally much larger than the confidence intervals for the means, which were discussed in the previous section. Confidence intervals for this model's projected variables are presented in Table 4.10. The intervals often exceed 20 percent and can be as

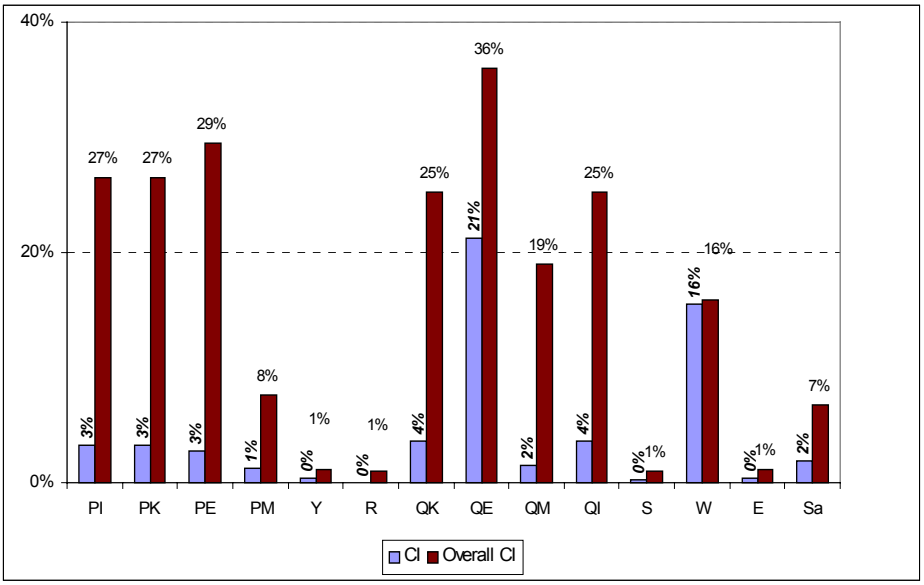
large as 30 percent. Confidence intervals for price and quantity variables are the are in excess of 20 percent. Few of the variables are determined very precisely.

Table 4.10: Confidence Intervals in Percent Deviation Accounting Residual Uncertainty

Variables	Confidence Interval
PI	26.30%
PK	26.31%
PE	29.31%
PM	7.56%
Y	1.14%
R	1.04%
QK	24.99%
QE	29.03%
QM	19.00%
QI	24.99%
S	0.95%
W	2.93%
E	1.14%
Sa	6.46%

This shows that residual uncertainty matters more than the parameter uncertainty for predicting the levels of the model's variables. Almost all of the variables are significantly affected by the residual uncertainty. Figure 4.6 illustrates the difference in the confidence intervals of the variables under parameter and residual uncertainty.

Figure 4.6: Comparison of Confidence Intervals



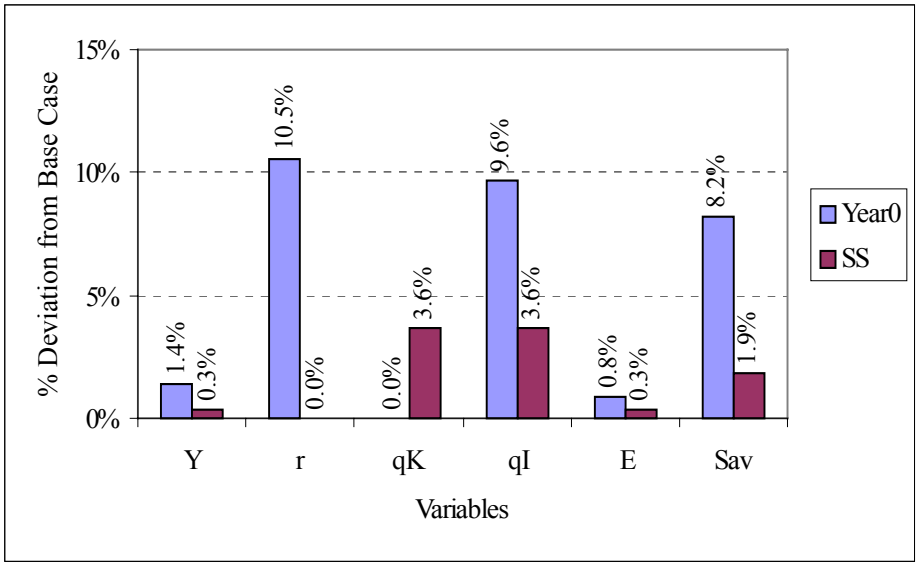
4.4.2 Intertemporal Results with Confidence Intervals[49]

The solution to the full intertemporal model was computed for 30 periods spaced two years apart. The first period represents the immediate short run and will be designated "year 0" below. The final period represents the steady state and will be designated "SS". The steady state conditions were imposed in the final period; in all other periods the evolution of the model was determined by its intertemporal equations. The initial capital stock was imposed in the first period. Confidence intervals for key variables were calculated at each point in time.[50]

[49] Only the confidence interval accounting for parameter uncertainty is presented.

[50] A total of 1440 equations are solved. With a grid spacing of two years for 30 periods shows the model for a total of 60 years into the future. The years are defined from year zero being the initial year and the last year as the steady state year. Note that solving intertemporal model is much more time consuming and difficult than solving a one period steady state model.

Figure 4.7: Comparison of Initial and the Steady State Confidence Intervals

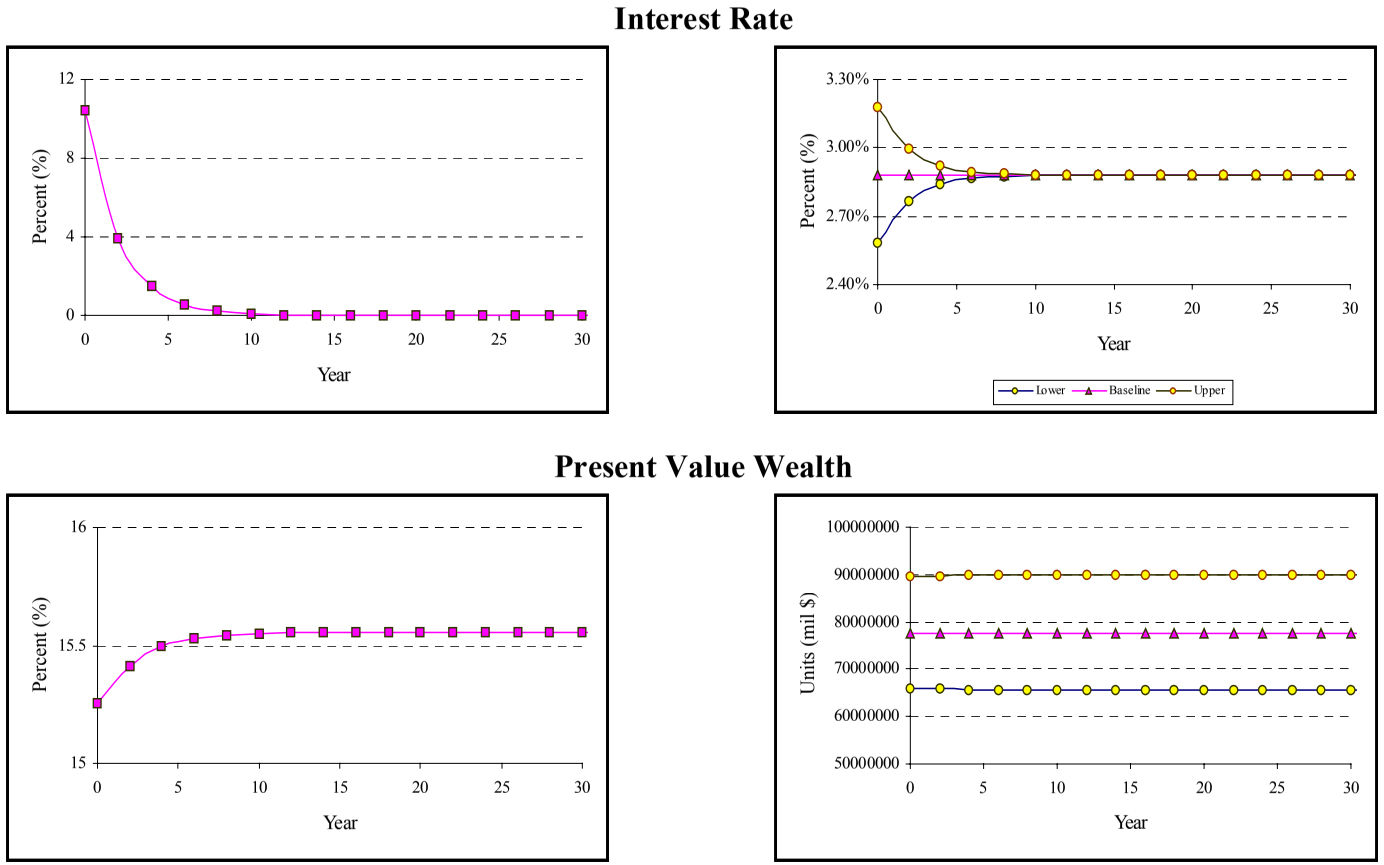


Initial and long run confidence intervals for some of the variables are illustrated in Figure 4.7. The intervals change markedly over time. The reason is interesting and intuitive: data ties down the model's state variable at the beginning of the simulation and theory ties down its costate at the end. In particular, the initial capital stock is known for certain, giving it a zero confidence interval year 0.[51] The interest rate in the steady state, on the other hand, is determined by theory and thus has a zero confidence interval in SS. Between the two periods, the confidence intervals evolve as expected: the capital stock becomes less and less certain, which the interest rate becomes more and more certain. The effects of these variables uncertainty drive the confidence intervals of the other variables. These effects can be seen clearly in Figure 4.8, which shows the intertemporal

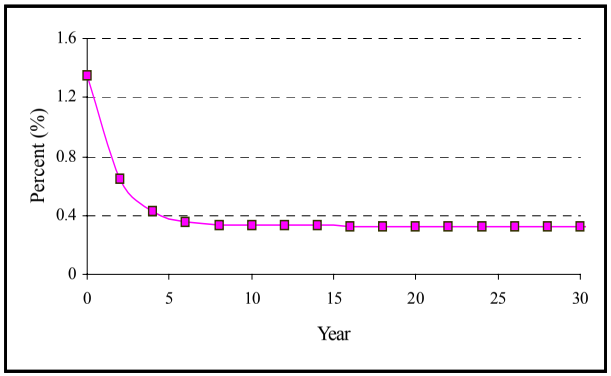
[51] This could be relaxed to incorporate uncertainty in the initial level.

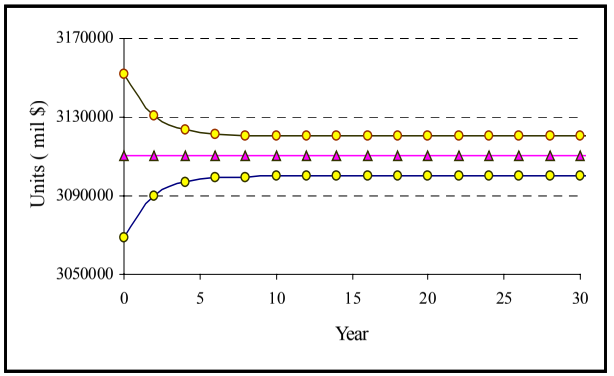
paths of the confidence intervals, both in levels and percentage deviations from the base case.

Figure 4.8: Intertemporal Confidence Intervals in Percent Deviations and Levels

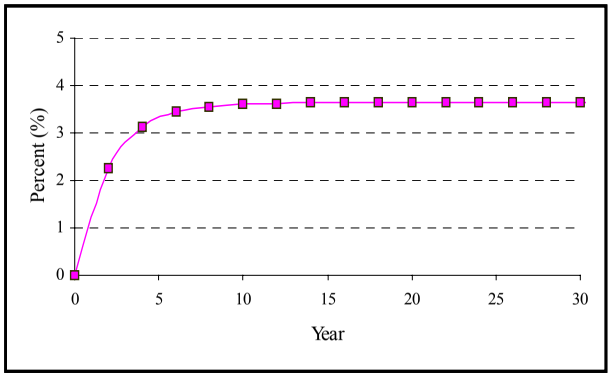


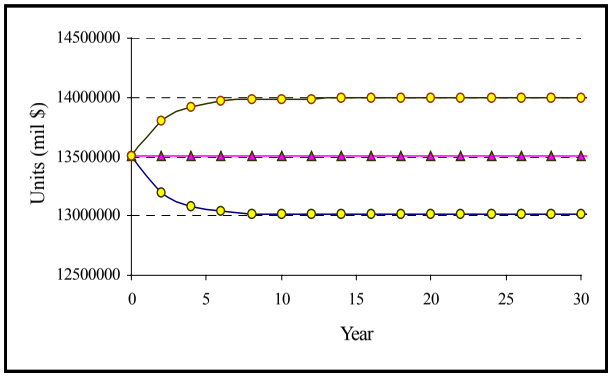
Income



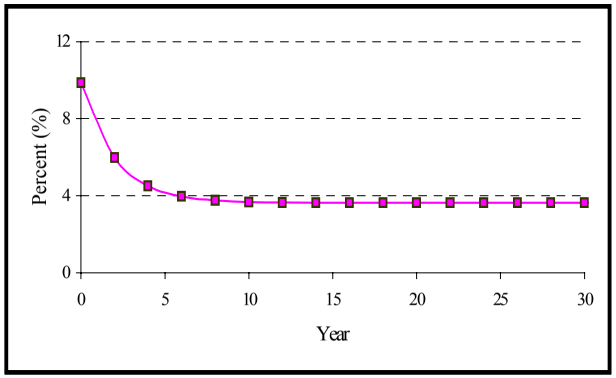


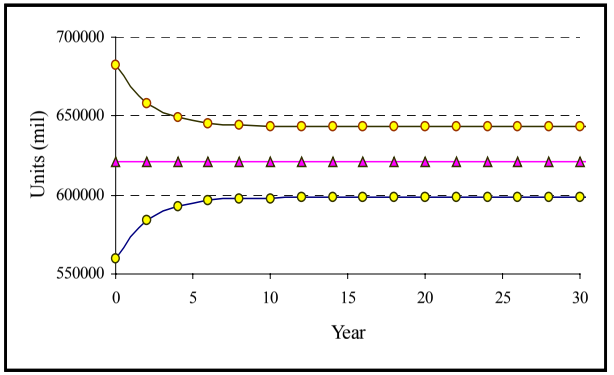
Capital



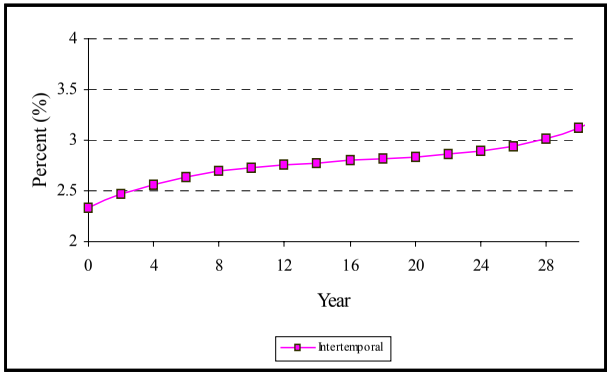


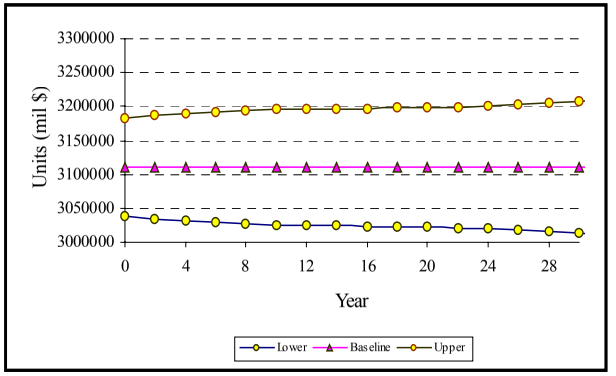
Investment





Savings





4.6 CHAPTER SUMMARY

In this chapter, an intertemporal econometric general equilibrium model of the United States was developed. The model featured three industries, a representative consumer, and a government. Fifteen producer parameters were jointly estimated by nonlinear least squares. Similarly, eight consumer parameters were estimated jointly by full information maximum likelihood. Parameter and residual covariance were recovered for confidence interval calculations.

The model was initially solved for the steady state and then extended to the full intertemporal solution. The confidence intervals for variables in the steady state model and the intertemporal model were calculated and compared. As in the previous chapter, accounting for residual uncertainty resulted in larger confidence intervals than accounting for parameter uncertainty alone.

This chapter also showed that the trajectory of the intertemporal confidence intervals varied widely for the variables. In other words, for some variables the confidence intervals were narrower at the initial period while others displayed larger confidence intervals. These differences arose because some of the variables are known precisely at beginning period while others are known precisely at the terminal period because of restrictions imposed by theory.

So far this dissertation has presented a method for computing confidence intervals for CGE models; has shown that the method is reliable (in the sense that it provides results very similar to Monte-Carlo simulation); and has demonstrated that it produces interesting results. The next step, which will be the subject of

Chapter 5, is to determine how confidence intervals can be integrated into policy analysis. The real gain from being able to compute a CGE model's confidence intervals is to be able to use that information to determine the precision of policy predictions, and to compute the confidence intervals of measures such as equivalent variation.

Chapter 5: Application to the Double Dividend Hypothesis

5.1 INTRODUCTION

This chapter demonstrates how confidence intervals can contribute to a real policy debate: a proposed environmental tax reform. The chapter provides a brief discussion of the "double dividend" issue and relates it to the model discussed in the previous chapter. The objective is not to present a detailed literature review of the double dividend hypothesis. Rather, the goal is to provide a new approach for understanding the source of the apparent contradictions between models used to examine the hypothesis, and to make an attempt to reconcile some of the results.

Market-based instruments have increasingly taken center stage in policy debates over controlling and tackling environmental problems. Instruments that raise substantial amounts of government revenue, including taxes and auctioned permits, have been advocated by some groups on the grounds that the revenue they generate can to be used to lower distorting taxes elsewhere in the economy or to reduce certain costs of the overall tax system in addition to improving the environment. This idea has become known in the literature as the "double dividend" hypothesis. Two versions of this hypothesis are extensively examined and discussed in the literature. Goulder (1994) refers to the claims as:[52]

[52] Although Goulder refers to yet another claim *Intermediate Form*, it is not discussed in this text because it not relevant, at this level, for the testing the hypothesis in this chapter.

> *"Weak Form*: By using revenues from the environmental tax to finance reductions in marginal rates of an existing distortionary tax, one achieves cost savings relative to the case where the tax revenues are returned to taxpayers in lump-sum fashion.
>
> *Strong Form*: The revenue-neutral substitutions of the environmental tax for typical representative distortionary taxes involve a zero or negative gross cost."

In another words, the weak form implies that using the revenue to reduce a distorting tax would be better than returning it as a lump sum rebate. However, the stronger version, stated differently, implies that taxing "dirty goods" and using the generated revenue to reduce the distorting tax would be better even before considering environmental benefit. The extreme version of this form can be viewed as: tighter regulations can be a "no regrets" policy that can be justified even if its environmental benefits are modest or impossible to quantify. The stronger form of the hypothesis has been far more controversial.

In the notation of Goulder (1994), if τ_E is the environmental tax, Δt_X is the cut in distortionary tax, and ΔT_L is the lump-sum tax reduction to achieve revenue-neutrality, then the weak form of the double dividend, in terms of the gross cost, *C*(.), for a given tax initiative, can be written as:

$$C(\tau_E, \Delta t_X) < C(\tau_E, \Delta T_L)$$

In another words, the gross cost is lower if the revenues are replaced through a cut in the distortionary tax than if the revenues are replaced through a lump-sum

rebate. However, the strong form asserts that swapping an environmental tax for a distortionary tax involves a negative gross cost. This form can be written as:

$$C(\tau_E, \Delta t_X) < 0$$

The weak form compares the costs of two different policies. However, the strong version is evaluated strictly by observing the sign of the change in welfare or costs, rather than the magnitude.

Many authors have studied the double dividend issue, both theoretically and empirically. Terkla (1984) estimates the potential efficiency gains from substituting revenues generated from effluent taxes on two major air pollutants for federal personal income tax and corporate income tax revenue. Lee and Misiolek (1986) examine the optimal size of a pollution tax, taking into account the efficiency gains caused by tax substitution. Results of these studies provide partial equilibrium support for the double dividend hypothesis.[53] However, Bovenberg and Mooij (1994) present a theoretical argument suggesting that general equilibrium effects can eliminate or even reverse the effect: that environmental taxes often exacerbate existing distortions, particularly in the labor market. Only a few studies have attempted to test the double dividend issue from a general equilibrium perspective. In a general equilibrium framework, Jorgenson and Wilcoxen (1993) measure the welfare impact of a carbon tax and Ballard and Medema (1993) focus on the efficiency of taxes in presence of certain types of

[53] Fullerton and Metcalf (1998) elaborate on the experiment design for both of these studies and the interpretation of their results.

externality. Both of these studies suggest that a double dividend can occur. Bovenberg and Goulder (1994), on the other hand, provide general equilibrium results that contradict the double dividend hypothesis. They find that the double dividend does not materialize when green taxes are substituted for traditional income taxes.[54]

Although theoretical contributions are significant and important in introducing and understanding the problem, the strong form of the double dividend hypothesis is basically an empirical issue. An empirical test of any form of the double dividend hypothesis depends on several factors. First, the initial state of the economy in which the reform of a double dividend is to be tested out matters. The results will depend on the magnitudes of the existing distortions (stemming from the initial tax rates): a double divided will be less likely if the green tax is added to a large existing tax; at the same time, it will be more likely if the tax to be reduced is highly distorting. Second, the magnitude of the reform under consideration could affect the results: the presence or absence of a double dividend for a small change in the tax system does not necessarily indicate whether a large change in tax rates would have a similar effect. Lastly, the structure of the model itself – its aggregation and functional forms – could be driving the results (see Fullerton and Metcalf (1998) for detailed discussion).

Tuladhar and Wilcoxen (1999) emphasize this idea through a stylized example of a potential green tax reform. In their simple illustration, the supply of capital is assumed to be elastic while the supply of labor is assumed to be

[54] Oates (1995) provides a concise update on the double dividend and caution that much work needs to be done before drawing policy conclusions.

perfectly inelastic. Under these assumptions, a tax swap that increases an energy tax and uses the revenue to reduce income taxes will produce a strong double dividend if the energy industry is more labor-intensive than rest of the economy. However, the reform would fail to generate a double dividend if energy is capital-intensive on average.

Since the double dividend hypothesis is an empirical question it must be tested empirically, using a model with econometrically estimated parameters and including confidence intervals for the results. A striking feature of all the studies done so far in the double dividend literature is that few studies use estimated parameters and none provide confidence intervals.[55] This chapter attempts to fill that gap by examining the strong form of the double dividend hypothesis using the econometric intertemporal general equilibrium model developed in the previous chapter. Unlike previous studies, the results will include confidence intervals.

In order to test the strong form of the hypothesis, a policy experiment is constructed that raises energy taxes and reduces capital taxes while keeping total tax revenue constant. The policy's mean equivalent variation (EV) and its standard error are calculated, which allows for statistical testing of the hypothesis. A positive value of the mean equivalent variation would indicate that the change leaves the household better off and would suggest that there is a strong double dividend. However, if the 95 percent confidence interval for the EV includes

[55] With reference to the global warming issue Poterba (1993) writes "While such uncertainty [uncertainty concerning the optimal carbon tax rate] is not unprecedented in tax policy discussions (what is the optimal tax rate on capital gains?), the sources of this uncertainty suggest promising directions for future research."

zero, then it would be impossible to reject the hypothesis that there is no double dividend. Similarly, the policy might generate a negative EV but if the confidence interval included zero, it would be impossible to reject the hypothesis that there is a double dividend. A firm conclusion for the reform simulated here can only be drawn if the confidence interval is either entirely positive (favoring the hypothesis) or entirely negative (refuting the hypothesis).

5.2 WELFARE ANALYSIS

5.2.1 Intertemporal Measure

Evaluating a policy requires assessing its effects on the household's level of well being. It is desirable from a policy maker's point of view to implement policies that improve welfare. In this section, an appropriate measure of intertemporal equivalent variation is constructed by means of the expenditure function. It is expressed as the change in wealth that would be equivalent to the policy in terms of its welfare impact.

In this model, the intertemporal analogue of the standard money-metric measure, Hicksian equivalent variation, is derived. The indirect utility function is obtained by substituting the optimal consumption levels into the direct utility function. Then, the optimal consumption levels are expressed in terms of wealth. The next step involves inverting the indirect utility function to get the intertemporal expenditure function. Intertemporal expenditure is expressed in terms of the total wealth required to attain any given level of intertemporal utility.

Intertemporal equivalent variation is given by the change in wealth needed to achieve the policy case utility at the base case prices.

Defining a new variable $\hat{W}(t)$ to be supernumerary wealth: the difference between actual wealth and the present value of the cost of the subsistence quantities:

$$\hat{W}(t) = W(t) - \zeta(p, \mu) \tag{5.1}$$

Using equation 5.1 allows 4.26 to be written as:

$$C^a(s) = \frac{\rho \hat{W}(t) e^{-\int_t^s (\rho - r(v))dv}}{\tilde{P}(s)} \tag{5.2}$$

Substituting the optimal intraperiod utility from equation 5.2 into the intertemporal utility function given in expression 4.12 allows the intertemporal indirect utility function to be written as:

$$V(t) = \int_t^\infty \ln\left(\frac{\rho \hat{W}(t) e^{-\int_t^s (\rho - r(v))dv}}{\tilde{P}(s)} \right) e^{-\rho(s-t)} ds \tag{5.3}$$

Taking the logarithm of the integrand in equation 5.3 reduces it to:

$$V(t) = \int_t^{\infty} \ln\left(\hat{W}(t)\right) e^{-\rho(s-t)} ds + \int_t^{\infty} \ln\left(\frac{e^{-\int_t^s (\rho - r(v))dv}}{\tilde{P}(s)}\right) e^{-\rho(s-t)} ds \tag{5.4}$$

Integrating the first term on the right-hand side of equation 5.4 and defining a long run interest rate R (discussed below) allows the expression to be written:

$$V(t) = \frac{1}{\rho} \ln\left(\hat{W}(t)\right) + \int_t^{\infty} \ln\left(\frac{\rho(R-\rho)(s-t)}{\tilde{P}(s)}\right) e^{-\rho(s-t)} ds \tag{5.5}$$

where for convenience the long run interest rate, R, from time t to s is defined to be:

$$R(t,s) = \frac{1}{s-t} \int_t^s r(v)dv$$

Expression 5.5 can be further simplified to the form:

$$V(t) = \frac{\ln\left(\hat{W}(t)\right)}{\rho} + \phi\left(\tilde{P}(\cdot), r(\cdot)\right) \tag{5.6}$$

where:

$$\phi\left(\tilde{P}(\cdot), r(\cdot)\right) = \int_t^{\infty} \ln\left(\frac{\rho(R-\rho)(s-t)}{\tilde{P}(s)}\right) e^{-\rho(s-t)} ds$$

Using equation 5.1 to eliminate $\hat{W}(t)$ from equation 5.6 and inverting it and recognizing that $\widetilde{P}$ is a function of p gives the intertemporal expenditure function:

$$W(p,V,r) = \zeta(p,\mu) + \exp\left(\rho V(t) - \rho\phi\left(\widetilde{P}(p), r\right)\right) \tag{5.7}$$

where p and r are the full trajectories of prices and interest rates from period t forward. W is thus a function of the price of consumption goods, the price of full consumption, the level of utility, and future interest rates respectively.

Finally, if $W_0(p_0, V_0, r_0)$ is the wealth associated with prices at the reference policy and $W_1(p_0, V_1, r_0)$ is the intertemporal expenditure required to attain the policy-case level of utility V_1 with at the reference policy prices, then intertemporal equivalent variation (EV) is given by the expression:

$$EV = W_1(p_0, V_1, r_0) - W_0(p_0, V_0, r_0) \tag{5.8}$$

where:

$$W_1(p_0, V_1, r_0) = \zeta(p_0, \mu) + \exp\left(\rho V_1 - \rho\phi\left(\widetilde{P}_0(p_0), r\right)\right) \tag{5.9}$$

and:

$$W_0(p_0, V_0, r_0) = \zeta(p_0, \mu) + \exp\left(\rho V_0 - \rho\phi\left(\widetilde{P}_0(p_0), r\right)\right) \tag{5.10}$$

It is convenient to express the equivalent variation as a percent of initial wealth because it is free of units and indicates the relative magnitude of the change in

welfare. Doing so produces expression 5.11, which will be used to evaluate the effectiveness of the policy:

$$\%EV = \frac{W_1(p_0, V_1, r_0) - W_0(p_0, V_0, r_0)}{W_0(p_0, V_0, r_0)} \tag{5.11}$$

5.2.2 Intraperiod Measure

Rewriting the expenditure function for the linear expenditure system gives:

$$E = \sum p_i(s)\mu_i + u\prod_i \left(\frac{p_i}{\alpha_i}\right)^{\alpha_i} \tag{5.12}$$

The intraperiod equivalent variation can be expressed as the change in expenditure at base case prices that would be equivalent to the policy in terms of its welfare impact. Thus it can be written as:

$$EV = \left(\sum p^0{}_i(s)\mu_i + u_1\prod_i \left(\frac{p^0{}_i}{\alpha_i}\right)^{\alpha_i}\right) - \left(\sum p^0{}_i(s)\mu_i + u_0\prod_i \left(\frac{p^0{}_i}{\alpha_i}\right)^{\alpha_i}\right) \tag{5.13}$$

Again, as done for the intertemporal equivalent variation, it is expressed as a percent of household expenditure:

$$\%EV = \frac{u_1 \prod_i \left(\frac{p^0{}_i}{\alpha_i} \right)^{\alpha_i} - u_0 \prod_i \left(\frac{p^0{}_i}{\alpha_i} \right)^{\alpha_i}}{u_0 \prod_i \left(\frac{p^0{}_i}{\alpha_i} \right)^{\alpha_i}} \text{ x } 100 \qquad (5.14)$$

Simplifying further, equation 5.14 can be written as the percentage change in utility between two periods. This is used to calculate the intraperiod equivalent variation.

$$\%EV = \frac{u_1 - u_0}{u_0} \text{ x } 100 \qquad (5.15)$$

5.3 EXPERIMENT DESIGN: THE EFFECT OF A SHIFT TOWARDS ENERGY TAXES

This section examines the double dividend hypothesis by simulating a shift in a tax policy. Initial tax rates on the primary inputs (capital and labor) are taken to be 28 percent and 10 percent respectively.[56] These two taxes combined form the preexisting distortions in the tax system of the economy. The "dirty good" of the economy, energy, is not initially taxed.[57]

The reform being considered is to increase the tax rate on energy to 10 percent from its initial value of zero and simultaneously reduce the tax rate on

[56] These numbers are approximate to the numbers used in Jorgenson and Wilcoxen (1990).

[57] This is assumed to simulate a case where there is no market based environmental regulation for the base year (1982) for the illustrative application. The tax policies before and after are clearly indicated. This clarity is important in understanding the policy results. As Fullerton and Metcalf (1997) point out: "The important point is that the evaluation must fully specify the policies already in place as well as the reform under consideration."

capital by exactly enough to leave the lump-sum subsidy unchanged.[58] The tax rate on labor is fixed at 28 percent. Table 5.1 summarizes the experiment.

Table 5.1: Tax Policy Experiment

Tax on	Base case	Policy
Capital	10%	Free
Labor	28%	28%
Energy	0%	10%

Tax revenues under the base case and the alternative policy are shown in Table 5.2. Total revenue collected from taxing energy amounts to about 23 billion dollars. The tax rate on capital is reduced just enough to balance the government's budget. This reduction is almost 2 percentage points, which reduces the capital tax rate to 8.17%. The revenue from capital taxes falls from 123 to 100 billion dollars.

Table 5.2: Revenue Generated Before and After the Policy Implementation
(in billions of 1982 dollars)

Revenue from taxing	Base Case	Policy
Capital	$123	$100
Labor	$716	$716
Energy	$0	$23

[58] The reform is focused on a tax on capital, rather than on labor, because labor supply is exogenously fixed. Thus, taxing labor would be equivalent to a lump-sum rebate. The tax rate are assumed to be inclusive of the income tax.

Tables 5.3 summarizes other key results for the short run (period 0) and the long run (the steady state), and presents percentage changes from the base case results. In is interesting to compare the results for the steady state, which shows the long run effect of the policy, with those for the complete transition path, which incorporates short and medium run effects as well. In the short run, the tax on energy results in a decrease in the relative price of capital services (due to the capital tax reduction), and increase in the relative price of materials and investment goods. This induces the household to substitute away from energy and into capital in the short run. Capital is fixed in the short run, so the increase in demand tends to push the rental price upward, offsetting part of the drop due to the change in taxes. However, capital accumulates in the long run, which drives the rental price down and the interest rate converges back to the time preference rate. In the long run, the household consumes more labor, materials, and capital services and less energy than the base case.

Table 5.3: Effects of the Simulation at the Aggregate Level

		Base Case	Short Run	% Change from Base Case	Long Run	% Change from Base Case
	Quantity of inputs					
Energy Industry	Capital	150.1	143.8	-4.24%	145.1	-3.35%
	Labor	61.8	58.7	-4.91%	58.7	-4.99%
	Energy	259.2	247.1	-4.66%	248.2	-4.24%
	Materials	66.5	63.1	-5.08%	63.2	-4.92%
Materials Industry	Capital	835.4	837.7	0.27%	840.5	0.61%
	Labor	1725.1	1726.4	0.07%	1727.2	0.12%
	Energy	169.8	166.1	-2.14%	166.5	-1.95%
	Materials	2119.4	2119.9	0.02%	2122.4	0.14%
Investment Demand	Capital	37.4	38.6	3.23%	38.4	2.70%
	Labor	110.7	113.4	2.44%	111.5	0.78%
	Energy	13.2	12.4	-5.99%	12.3	-6.98%
	Materials	523.1	534.8	2.25%	527.7	0.87%
Consumption	Capital	662.0	664.9	0.44%	674.8	1.94%
	Labor	359.4	358.5	-0.25%	359.6	0.05%
	Energy	199.5	186.1	-6.71%	187.5	-6.01%
	Materials	2086.3	2078.5	-0.37%	2088.763	0.12%
Purchase Price of	Energy	$0.6453	$0.7076	9.65%	$0.7033	8.98%
	Materials	$1.0009	$1.0029	0.20%	$0.9999	-0.10%
	Capital Services	$0.7345	$0.7284	-0.82%	$0.7198	-2.00%
Capital Stock		13500	13500.5	0.00%	13612.14	0.83%
Investment		621.0	634.8	2.21%	626.2	0.83%

Note: Unless otherwise stated, the units are in 1982 million model units.

The overall welfare effect of the policy is small and negative: the intertemporal equivalent variation is -0.038%, or a drop equivalent to about four hundredths of a percent of initial wealth. The increase in the tax on energy apparently causes intertemporal welfare losses that more than offset the gains due to the decrease in the tax on capital. Judging from the mean results alone, the tax policy appears to refute the strong form of the double dividend hypothesis. However, it remains to be seen whether this result is statistically significant: that is, whether the model's confidence intervals are narrow enough to reject the double dividend hypothesis for the simulated reform.

To illustrate the importance of dependence of a double dividend effect on the initial state of the economy, another experiment was performed in which the initial tax rate on capital was set to 15% instead of 10%, increasing the tax distortion. The reform simulated was the same as before: to increase the tax rate on energy from zero to 10 percent and simultaneously reduce the tax rate on capital by exactly enough to leave the lump-sum subsidy unchanged. The tax rate on labor was fixed at 28 percent. The overall welfare effect of the policy was positive: the intertemporal equivalent variation was 0.044.

5.4 INTERTEMPORAL EQUIVALENT VARIATION CONFIDENCE INTERVALS

The overall confidence interval (incorporating both parameter and residual uncertainty) for the policy's intertemporal equivalent variation runs from –0.67 to 0.59 percent.[59] Note that the overall confidence interval includes zero and that the

[59] Calculation of confidence intervals of equivalent variation involves comparison of two simulations, the policy model and the base case model, then substituting the expression 5.11.

mean value, -0.038, is very close to zero relative to the width of the interval. This implies that the uncertainties in the parameter estimates and in the overall fit of the regression equations are large enough that it is impossible to reject the double dividend hypothesis at the 95% level. In other words, there are combinations of parameter estimates, consistent with the estimated variance-covariance matrices, that would result in a positive equivalent variation. For example, if capital consumption subsistence parameter is assumed to take a value that is two standard errors above its mean value then the equivalent variation would be positive 0.35 percent. Similarly, if the same subsistence parameter is fixed at zero (collapsing the utility function to Cobb-Douglas), then the intertemporal equivalent variation would be 0.08 percent, again a positive value. In short, it is easily possible for the equivalent variation's sign to be reversed if the true values of imprecisely estimated parameters differed modestly (that is, within their own confidence intervals) from their estimates. Thus, the overall confidence interval indicates that the hypothesis that there is no double dividend cannot be rejected.[60]

Figure 5.1 compares the overall confidence interval for the equivalent variation with the confidence intervals generated by considering parameter uncertainty and the estimation residuals separately. The confidence interval due

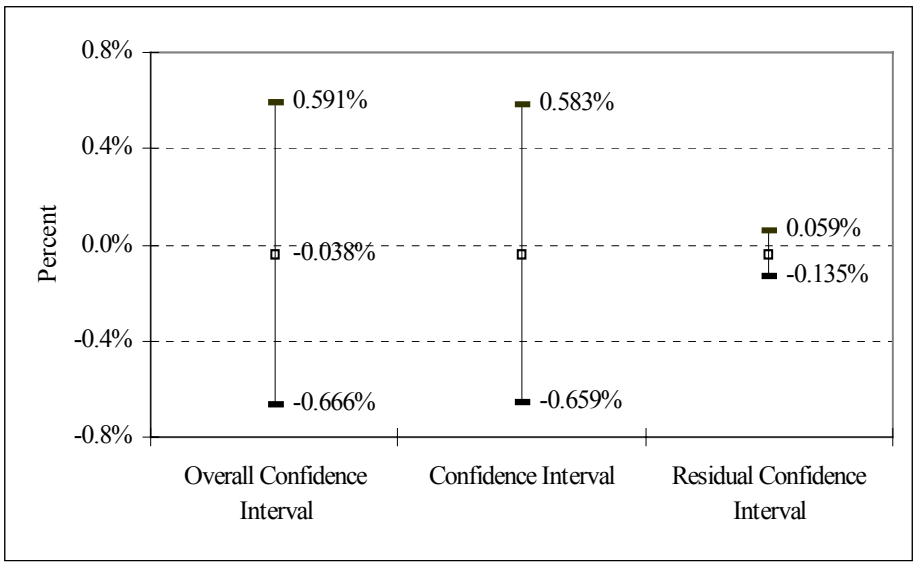
---

Computation can be simplified by evaluating both models simultaneously by stacking the policy model on top of the base case model. However, this doubles the number of equations to be solved and it is time-consuming to evaluate the Jacobian. The stacked intertemporal model consists of 112 equations. If the steady state is imposed after 60 years, then the full model involves 6720 equations. The biggest challenge in calculating the confidence intervals for this model is to convert the GEMPACK output to a text that is readable in GAUSS.

[60] However, the confidence interval for the intertemporal equivalent variation for the policy simulation with the initial tax rate on capital set to 15% runs from –0.196 to 0.284 percent. The interval includes zero, thus for the simulation, the hypothesis that there is a double dividend cannot be accepted.

to imprecision in the parameters is nearly as large as the overall confidence interval: from -0.66 to 0.58. The confidence interval accounting for unexplained variation is much narrower and ranges from only negative 0.13 to positive 0.06.

Figure 5.1: Comparison of Intertemporal Equivalent Variation



The similarity in magnitude between the confidence interval (parameter uncertainty only) and the overall confidence interval suggests that imprecision in the parameter estimates is the most important determinant of imprecision in the equivalent variation. The confidence interval due to unexplained variation is much smaller. This is an interesting and important result: even though the unexplained variation in the regression equations is the most important source of

imprecision in the overall levels of the model's variables, it is not an important source of imprecision in the equivalent variation calculation.[61]

5.5 UNDERSTANDING THE CONFIDENCE INTERVAL RESULTS

To understand the intertemporal equivalent variation result in more detail, it is helpful to calculate the intraperiod equivalent variations and their confidence intervals for key periods. The intraperiod EVs are calculated using equation 5.15 and only reflect the second stage of the household's optimization problem: the allocation of expenditure across commodities. They show the amount of expenditure needed to compensate for the policy's effect on that period alone, holding the rest of the trajectory at the base case.

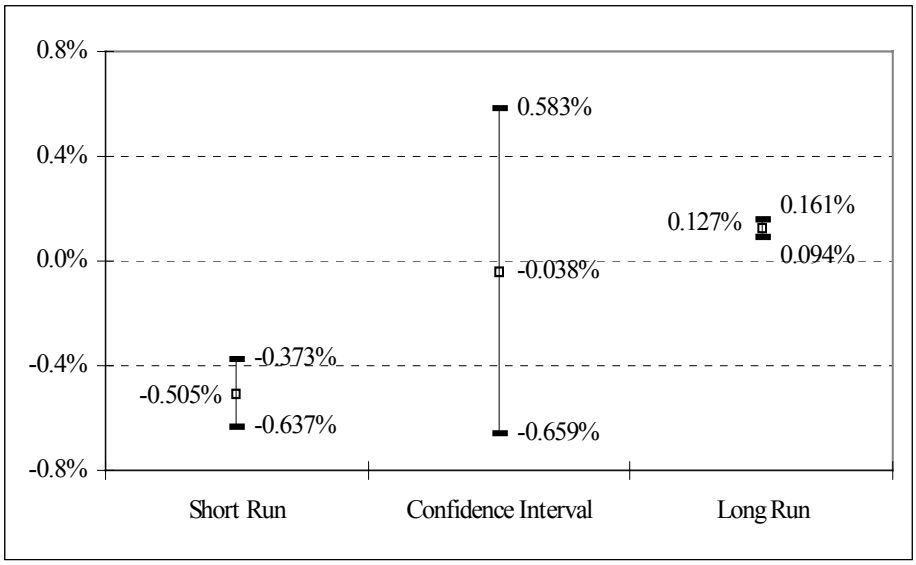
Two periods are of particular interest: the immediate short run (period 0) and the very long run (the steady state). Intraperiod equivalent variation confidence intervals for those periods are shown in Figure 5.2, along with the overall intertemporal confidence interval (expressed as a percentage of wealth).

The short and long run confidence intervals reveal how the period-by-period effects of the policy will play out over time. The results are distinctly different from the intertemporal equivalent variation. The mean of the short run intraperiod equivalent variation is -0.51 percent and the confidence interval runs from -0.64 to -0.37. The long run intraperiod equivalent variation, on the other hand, ranges from 0.09 to 0.16 with a mean value of 0.127 percent. Unlike the

[61] In a sense it is surprising that the residuals matter at all since they are held constant between the base case and the policy experiment. However, by changing the relative levels of variables in the base case (including wealth), they have a small effect.

intertemporal confidence interval, however, *neither* intraperiod confidence interval includes zero. The short run equivalent variation is exclusively negative and the long run equivalent variation is exclusively positive. This provides an immediate explanation for some of the variation appearing in the double dividend literature: models with a short-run emphasis would tend to reject the hypothesis while those with a longer run emphasis would tend to confirm it.

Figure 5.2: Equivalent Variation Confidence Intervals



The short and long run intraperiod equivalent variations differ so sharply for a very intuitive reason. In the long run, the tax shift is beneficial because it raises the capital stock and increases both the economy's output and household consumption. In the short run, however, households have difficulty substituting away from energy toward capital because capital is fixed in the short run. The

policy looks, in other words, like the imposition of a distorting tax on energy accompanied by an essentially lump-sum rebate. As a result, the short-run welfare effects show exactly the expected result: welfare falls. However, in the long run, the household is better off because the capital stock increases, allowing the household to substitute capital for energy.

It is also interesting to note that the intraperiod confidence interval in the long run is narrower than that for the short run. The reason is that the theoretical conditions imposed on the steady state, particularly that the interest rate must equal the time preference rate, are quite strong. This can be seen in the panels of Figure 5.3, which show the trajectories of key variables over time. (The figure also includes results for two alternative simulations, which will be explained below.)

Although the long run and the short run results are striking, neither should be mistaken for the model's true confidence interval. Overall conclusions about the double dividend hypothesis should be based on the intertemporal confidence interval because it accounts for the variations over all the time periods.

To examine the source of the results in more detail, two additional simulations were carried out. In the first of these, SIMK, the tax rate on capital income was reduced from its base case value of 10 percent to 8.17 percent, which was the result achieved by the energy-capital tax experiment. In the second experiment, SIME, the tax rate on energy was set to 10 percent (as in the energy-capital experiment) but the tax rate on capital was left at its base case value. In both experiments changes in the lump-sum tax were used to keep the

government's budget in balance. Roughly speaking, the two experiments decompose the energy-capital tax swap: SIMK looks at the benefits of cutting the capital tax and SIME looks at the costs of raising the energy tax.[62] The structure of the simulations is summarized in Table 5.4 and key results are shown in Table 5.5 and Table 5.6.

Table 5.4: Simulations Proposed to Evaluate Individual Tax Effects

Tax on inputs	Simulations			
	tax on capital SIMK		tax on energy SIME	
	Base case	Policy	Base case	Policy
Capital	10%	8.17%	10%	10%
Labor	28%	28%	28%	28%
Energy	0%	0%	0%	10%

The intertemporal equivalent variation for SIME is -0.036 percent and 0.025 percent for SIMK. A feature of these results is striking: both are very small, especially compared to the confidence intervals discussed above.

The tax on energy in SIME causes the household to substitute away from energy to other inputs rather quickly in the initial years, as shown in Figures 5.3g-5.3j. Under SIME, capital, labor, and materials reach a new steady state that is above the base case level, but energy consumption is much lower than the base case level. Under SIMK, however, the adjustment to the new steady state is slower than the simulation SIME because of the capital constraint in the short run.

[62] SIMK and SIME are suggestive but not a true decomposition of the original experiment because the capital tax rate is not constant in the original experiment.

SIMK induces a decrease not only in the rental price of capital (see Figure 5.3c) but also in the prices of other inputs, Figure 5.3a-5.3b. In the long run, the household consumes more capital, energy, and materials, but less labor.

The gain in welfare in the simulation SIMK due to an increase in capital, energy, and materials outweighs the loss in welfare due to a decrease in labor. Likewise, for simulation SIME the increase in capital, labor, and materials outweigh the loss due to decrease in energy. There is net positive gain in welfare from the SIMK than the SIME simulation. This indicates that the distortion due to tax on energy is more pronounced than the capital tax distortion when analyzed independently. The aggregate effects for SIMK and SIME are shown in Table 5.6 and Table 5.7, respectively. The transition paths of some key variables for the simulations are presented in Figures 5.3a - 5.3o.

Table 5.5: Effects of SIMK Simulation at the Aggregate Level

		Base Case	Short Run	% Change from Base Case	Long Run	% Change from Base Case
	Quantity of inputs					
Energy Industry	Capital	150.1	150.3	0.14%	152.5	1.61%
	Labor	61.8	61.4	-0.53%	61.4	-0.64%
	Energy	259.2	258.7	-0.19%	260.5	0.51%
	Materials	66.5	66.3	-0.36%	66.4	-0.09%
Materials Industry	Capital	835.4	837.1	0.20%	841.4	0.72%
	Labor	1725.1	1725.3	0.01%	1726.5	0.08%
	Energy	169.8	170.0	0.10%	170.5	0.41%
	Materials	2119.4	2120.6	0.05%	2124.4	0.24%
Investment Demand	Capital	37.4	38.8	3.95%	38.6	3.20%
	Labor	110.7	114.2	3.19%	111.4	0.69%
	Energy	13.2	13.7	3.58%	13.5	1.97%
	Materials	523.1	540.8	3.38%	529.9	1.31%
Consumption	Capital	662.0	658.6	-0.50%	673.9	1.80%
	Labor	359.4	356.1	-0.93%	357.7	-0.48%
	Energy	199.5	198.1	-0.68%	200.4	0.49%
	Materials	2086.3	2070.3	-0.77%	2085.9	-0.02%
Purchase Price of	Energy	$0.6453	$0.6427	-0.40%	$0.6366	-1.34%
	Materials	$1.0009	$0.9989	-0.20%	$0.9943	-0.65%
	Capital Services	$0.7345	$0.7288	-0.78%	$0.7154	-2.60%
Capital Stock		13500.5	13500.5	0.00%	13673.2	1.28%
Investment		621.0	642.0	3.37%	629.0	1.28%

Note: Unless otherwise stated, the units are in 1982 million model units.

Table 5.6: Effects of SIME Simulation at the Aggregate Level

		Base Case	Short Run	% Change from Base Case	Long Run	% Change from Base Case
	Quantity of inputs					
Energy Industry	Capital	150.1	143.4	-4.50%	142.9	-4.84%
	Labor	61.8	59.3	-4.03%	59.0	-4.38%
	Energy	259.2	247.9	-4.36%	247.0	-4.71%
	Materials	66.5	63.5	-4.48%	63.3	-4.83%
Materials Industry	Capital	835.4	834.7	-0.08%	834.6	-0.10%
	Labor	1725.1	1726.2	0.06%	1725.8	0.04%
	Energy	169.8	165.8	-2.33%	165.8	-2.34%
	Materials	2119.4	2117.8	-0.07%	2117.4	-0.09%
Investment Demand	Capital	37.4	35.9	-3.81%	37.2	-0.42%
	Labor	110.7	107.0	-3.29%	110.8	0.11%
	Energy	13.2	11.7	-11.84%	12.1	-8.74%
	Materials	523.1	503.3	-3.79%	521.0	-0.41%
Consumption	Capital	662.0	670.8	1.33%	663.1	0.17%
	Labor	359.4	364.5	1.42%	361.3	0.53%
	Energy	199.5	188.3	-5.60%	186.6	-6.45%
	Materials	2086.3	2107.1	1.00%	2089.1	0.14%
Purchase Price of	Energy	$0.6453	$0.71	10.44%	$0.71	10.44%
	Materials	$1.0009	$1.01	0.55%	$1.01	0.55%
	Capital Services	$0.7345	$0.74	0.57%	$0.74	0.57%
Capital Stock		13500	13500.52	0.00%	13443.36	-0.42%
Investment		621.0	597.4	-3.81%	618.4	-0.42%

Note: Unless otherwise stated, the units are in 1982 million model units.

5.6 CHAPTER SUMMARY

This chapter has demonstrated the importance of computing confidence intervals for CGE results by examining a current topic in the environmental literature: the strong form of the double dividend hypothesis. The specific policy examined, as an illustration, was an increase in the tax on energy accompanied by a decrease in the capital tax rate. In particular, the tax rate on energy was increased to ten percent and the model was solved for an accompanying revenue-neutral reduction in capital taxes.

The mean value of the intertemporal equivalent variation was found to be negative but the overall confidence interval included zero and was very broad. Thus, the hypothesis that there is no double dividend cannot be rejected for the proposed reform. The confidence interval of the mean value was also calculated, in order to determine whether the imprecision of the equivalent variation was due to uncertainty in the parameter estimates or to the residuals in the regression equations. It was found to be nearly identical to the overall confidence interval. Thus, the unexplained variation in the estimating equations was relatively unimportant in determining the confidence interval for the equivalent variation; all that really matters is the variance-covariance matrix for the parameter estimates.

Figure 5.7: Transition Paths for Different Tax Reforms

Figure 5.3a: Purchase Price of Energy ($)

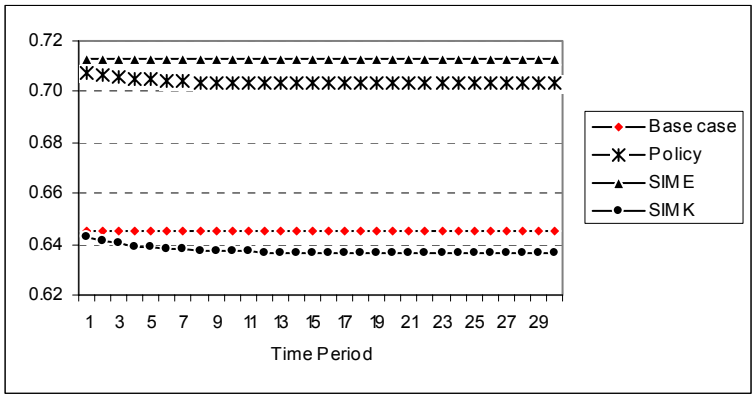


Figure 5.3b: Purchase Price of Materials ($)

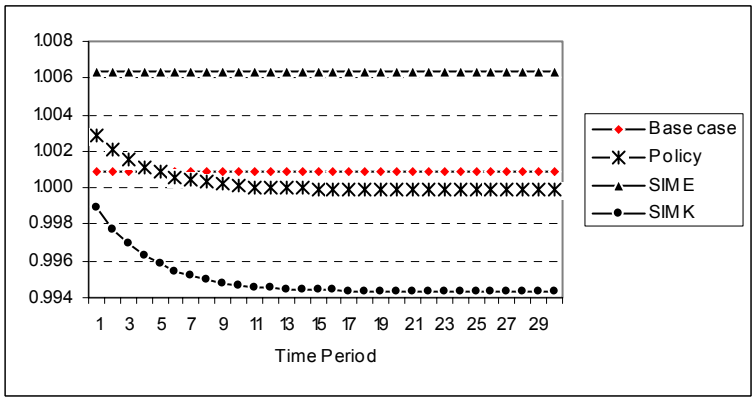


Figure 5.3c: Rental Price of Capital ($)

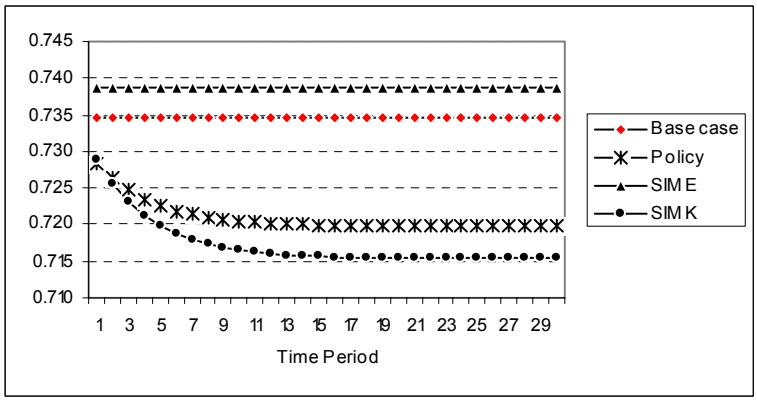


Figure 5.3d: Income (bil. $)

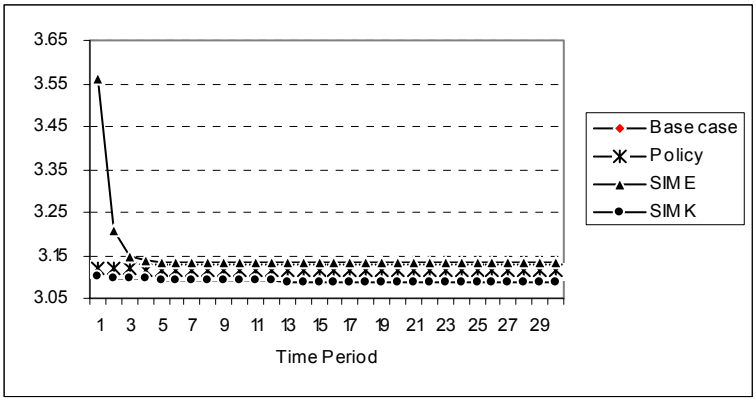


Figure 5.3e: Expenditure (bil. $)

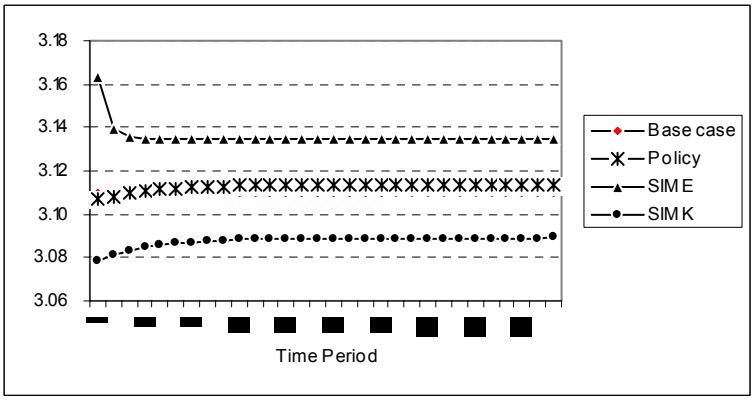


Figure 5.3f: Present Value Wealth (tril $)

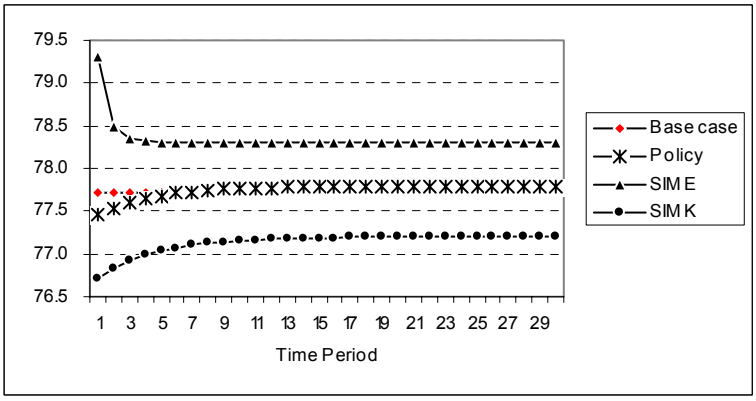


Figure 5.3g: Household Consumption of Energy (tril. units)

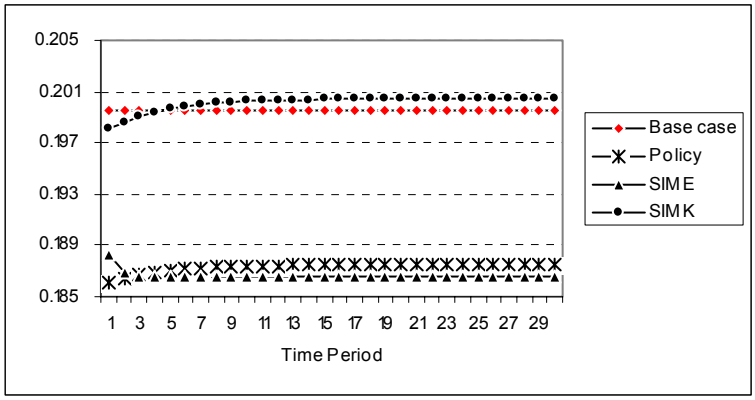


Figure 5.3h: Household Consumption of Materials (bil. units)

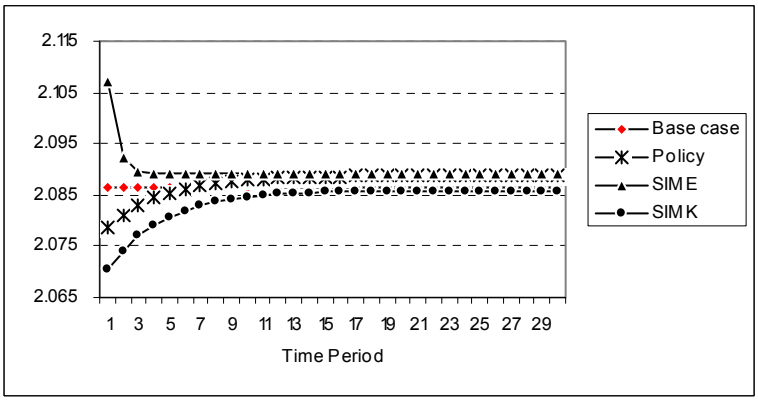


Figure 5.3i: Household Consumption of Capital Services (bil. units)

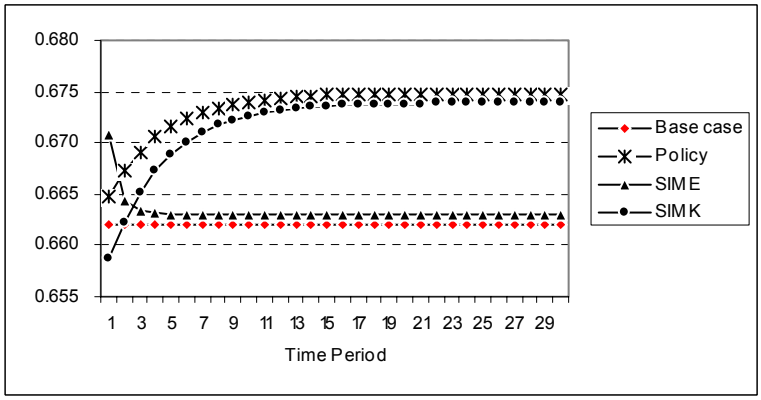


Figure 5.3j: Household Consumption of Labor services (10^7 hours)

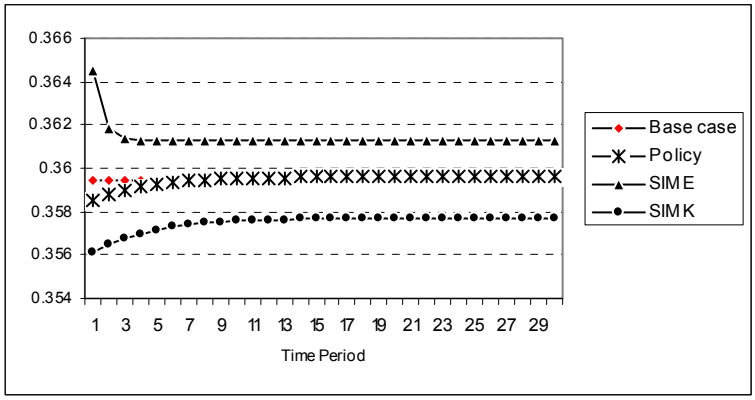


Figure 5.3k: Capital Stock (bil. units)

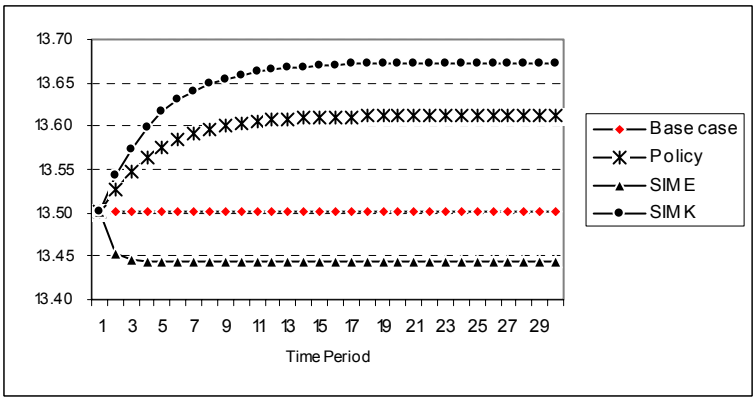


Figure 5.3l: Total Energy Demanded (bil. units)

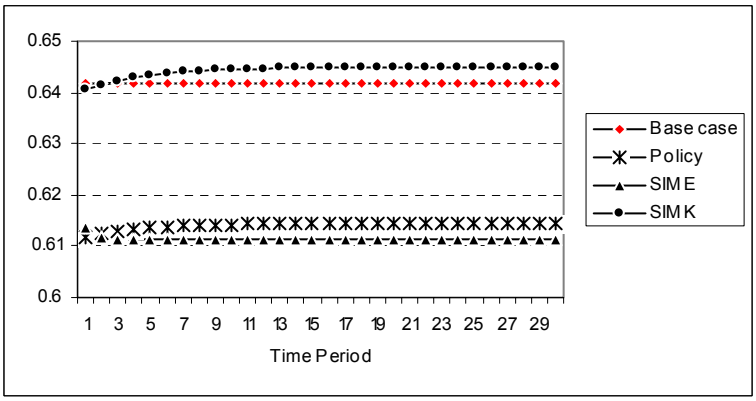


Figure 5.3m: Total Materials Demanded (bil. units)

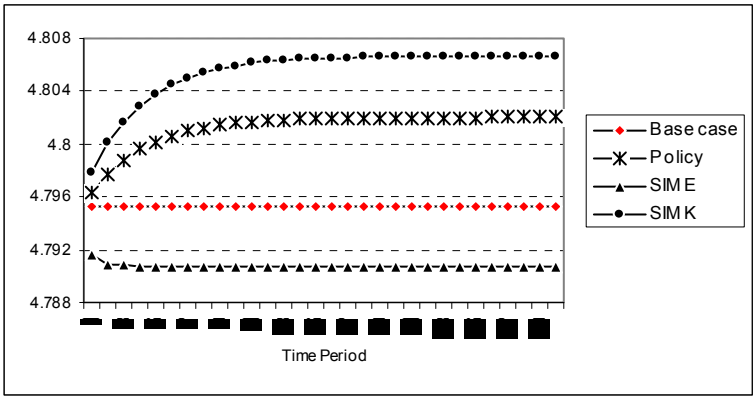


Figure 5.3n: Interest Rate

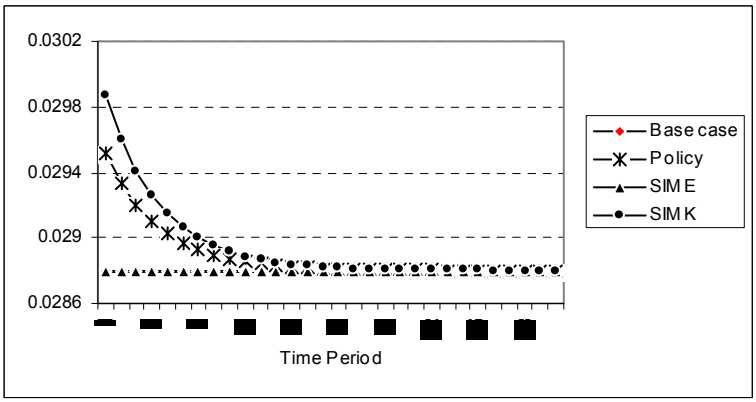
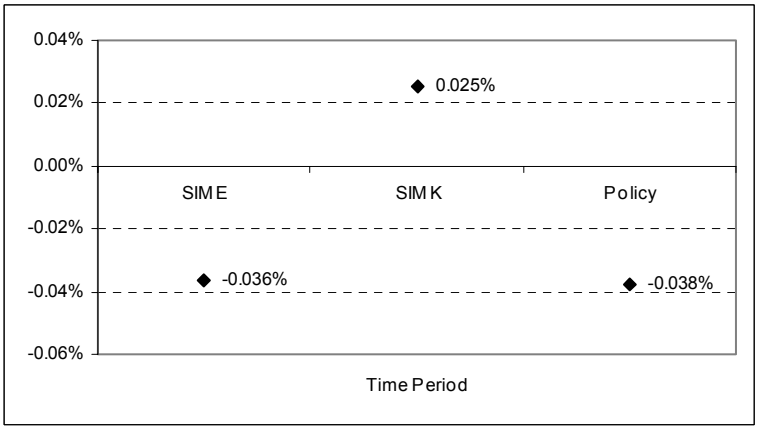


Figure 5.3o: Intertemporal Equivalent Variation (%)



Chapter 6: Conclusions and Caveats

6.1 CONCLUSIONS

Computational general equilibrium modeling plays an increasingly important role in the analysis of major economic policies. For example, CGE models have been used to analyze environmental policy, tax policy and trade reform. However, CGE models involve many, many parameters that are not known precisely. Moreover, the econometric equations used in CGE models do not explain all of the variation in historical data. These two features imply that CGE simulation results should be accompanied by standard errors or confidence intervals whenever they are reported. For example, a CGE analysis of a proposed tax reform should not simply report that it would raise welfare by a given percent; instead, it should report that welfare would rise by the given percent plus or minus a two standard-error range. Without such information, policy makers have no way to judge the degree of certainty that should be associated with a model's results. The degree of certainty could be a critical attribute of a policy: a tax reform that raises welfare by 2 percent plus or minus 5 percent is probably much less attractive than one that raises welfare by 1 percent plus or minus 0.5 percent. Despite the evident importance and usefulness of confidence intervals, they are virtually never reported for CGE results. This dissertation attempts to rectify that problem. Its key insights and objectives can be broadly stated as follows:

- A natural way to quantify a model's parameters is to estimate them econometrically. The variance-covariance matrix for the estimates can then be used to quantify the precision of the model results. Under this approach, CGE modeling approach can be regarded as an econometric exercise and all results should be accompanied by standard errors or confidence intervals;
- The principle contribution of this research is to propose and then demonstrate a systematic approach to computing confidence intervals for CGE models.
- Since formal confidence intervals have not previously been used to quantify the precision of CGE results, an important question examined by this dissertation is how confidence intervals compare to the sensitivity analysis that is sometimes used in CGE modeling;
- Finally, the dissertation provides a practical example of the use of CGE confidence intervals by examining an important and controversial topic in the recent literature: the strong form of the double dividend hypothesis.

The introduction explores the question of why CGE results have not in the past been presented with confidence intervals. One possible reason is that some modelers have failed to realize that CGE modeling is an econometric exercise, and thus do not realize the need to report confidence intervals. This seems especially likely with models that use calibrated parameters. Other researchers, however, have been concerned about quantifying and reporting the precision of their results but have lacked a rigorous methodology for doing so. Hence, at best, sensitivity analysis is performed. Although sensitivity analysis is arguably an

improvement over failing to test the robustness of a model's results at all, it is a very imperfect measure. Further, it is especially difficult to carry out sensitivity analysis on models with large numbers of parameters.

Like the proposed methodology, Kim (2003) is a step in the right direction that also recognizes the need to model uncertainty. The method proposed here, however, differs from Kim's work in the modeling approach used as well as the focus of the analysis. First, in the control theory approach the goal is to stabilize the economy in the face of an external shock. Thus, the focus is on the adjustment process of the control variables (exogenous policy variables). The proposed method, on the other hand, focuses on the magnitude of uncertainty associated with the endogenous variables (including the target variables of concern to the policy maker). It provides a clear and direct measurement of the precision of the results. Second, an important aspect of the optimal control with learning approach is that it updates the covariance matrices associated with the state (endogenous) variables and the control variables as new information arrives. However, there is no "learning" under the proposed CGE method because no new information becomes available at the time the simulation is run; that is, the covariance matrices remain constant for the simulation period.

In Chapter 2, a simple partial equilibrium model for total revenue was developed to illustrate the issue. The model's parameters were econometrically estimated and the overall confidence interval (which includes uncertainty in both the parameter estimates and the residuals) was calculated. In addition, a confidence interval for the mean was calculated, reflecting uncertainty in the

parameter estimates but excluding the unexplained variation in the econometric equations used for estimating the model's parameters. These confidence intervals are compared with the results of sensitivity analysis. The chapter concludes that sensitivity analysis results should be interpreted with caution because ad hoc selection of the parameters to test, and the amounts by which to perturb them, could lead to large discrepancies between the results of sensitivity analysis and the model's true confidence intervals.

Chapter 3 develops a small static econometric general equilibrium model of the U.S. economy, with a representative producer and a consumer. It presents the approach for calculating CGE confidence intervals in detail and presents numerical results for the model. Some of the findings are that: (1) uncertainty about the true values of the parameters affect some variables much more than others; and (2) uncertainty stemming from the estimation residuals is a major contributor to the confidence intervals for the results. The chapter also compares the confidence intervals to the results of a Monte-Carlo simulation and shows that the two methods yield very similar results. In addition, the confidence intervals are compared to the results from sensitivity analysis. Overall, Chapter 3 provides an econometric approach for computing CGE confidence intervals that can be performed relatively easily and routinely.

This approach is extended in Chapter 4 to a much more realistic intertemporal model of the U.S. economy. The model consists of three industries, a representative consumer and the government. Each industry uses energy, materials, labor, and capital services as inputs. The government collects taxes on

the inputs and returns the revenue as a lump-sum rebate. The producer and the consumer models are each estimated using a time-series data base. In total, twenty-three parameters are estimated. The model is solved for both the steady state equilibrium and the full intertemporal transition path. Confidence intervals are calculated for both types of solution.

Chapter 5 extends chapter 4 by developing a method for calculating the confidence intervals associated with comparisons between policies, such as an equivalent variation. As an example, it applies the method to evaluating the welfare effects of a possible tax reform: increasing energy taxes and using the revenue to reduce income taxes. In particular, one version of the strong form of double dividend hypothesis is tested. The principle finding is that the confidence interval for the policy's equivalent variation is very wide and includes zero. It is thus not possible (given current data) to conclude that there is or is not a strong double dividend. In addition, a detailed analysis of the results provides an explanation for why different modelers have reached sharply different conclusions.

An additional result from Chapter 5 is that unexplained variation in the model's estimating equations makes only a minimal contribution to the confidence intervals for an equivalent variation. Uncertainty about the true values of the model's parameters is far more important.

6.2 CAVEATS AND LIMITATIONS

The approach presented here is a tractable method that, if widely used, would sharply improve the quality of policy analysis carried out with CGE models. There are two minor caveats that must be kept in mind but that do not significantly detract from the methodology. First, the confidence interval calculation is done using a linearized version of the model and it is assumed that higher-order terms in the model's Taylor series expansion can safely be ignored. There is abundant precedent for using such an approximation since nonlinear estimation packages routinely use linearization for computing standard errors. Moreover, the results of the method are very similar to those obtained via Monte-Carlo simulation using the full nonlinear form of the model. Nonetheless, the confidence intervals should formally be regarded as approximations.

A second relatively minor caveat is that the real world models used for policy analysis are usually too large and complex for large subsections of the model to be estimated simultaneously. The number of parameters can run into the hundreds. Instead of estimating the model's entire production side at once, it is often necessary to estimate each industry separately, or even to estimate individual industries as a series of separate tiers. Estimating the model as a series of separate units has the effect of restriction the covariance matrices of the parameter estimates and the residuals. The important thing to note here, however, is that this is a limitation of the econometrics underlying a given model, not of the method for calculating confidence intervals. That is, the method can be applied

successfully even when it is impossible to estimate the model with complete flexibility.

In addition to these minor points, there is one major issue that remains to be resolved: developing an appropriate method for dealing with models that use a large number of parameters whose values have been imposed or calibrated, rather than being estimated. In an ordinary regression, restricting a parameter or imposing a function form will generally reduce the fit of a model and increase its unexplained variance. Imposing too much structure on the model, in other words, has an easily observed and widely understood consequence. The same would be true for a CGE model if a parameter in one of the estimating equations were restricted: the imprecision in the model's results arising from the residuals would rise and the confidence intervals for the model's variables would broaden.

The conceptual problem arises when many parameters are imposed without any attempt being made to calculate the consequence on the overall fit of the model. Each parameter imposed by fiat will reduce the rank of the variance-covariance matrix for the parameters and will thus be likely to narrow the model's reported confidence intervals. In the limiting case, consider a model with many calibrated parameters but only one estimated parameter. Such a model would have a single parameter variance and a single residual variance. Its confidence intervals would, as a result, be far narrower than their true values. A mechanism must be developed, therefore, that would distinguish confidence intervals built using extensive data collection and estimation from those with much weaker empirical underpinnings.

Appendix A: Description of the Data Used

The data used in this thesis is based on the data used in Wilcoxen (1988), which is in the form of inter-industry tables. These annual time series data on inter industry transaction tables for the US economy for the period 1947-1985 was constructed based on the input-output tables for the US published by the Bureau of Economic Analysis (1984).

The inter-industry table consists of 35 industries. In Figure A1, the columns of U represent the 35 industries and the rows represent the domestic commodities. The description of the industries is outlined in Table A1. Beside the commodities, industries also use three primary factors: non-competing imports (N), capital services (K), and labor (L). Other inputs are referred as rest of the world (R). The commodities and some of the primary factors are also demanded by the final demand sector: consumption (C), investment (I), government spending (G), imports (I), exports (M). The section V represents value added to each of the final demand categories.[63]

The data used in this thesis for the partial equilibrium model (Chapter 2), small static CGE model (Chapter 3), and the intertemporal CGE model (Chapter 4 and 5) is aggregated from 39 tables. The aggregation process for each of the data sets is explained below.

[63] Further details on the data is provided in Appendix C in Wilcoxen (1988).

Figure A1: Original Inter-industry Table

<table>
<tr><td>U
Commodity x industry
35 x 35</td><td>C</td><td>I</td><td>G</td><td>X</td><td>M</td></tr>
<tr><td>N</td><td colspan="5" rowspan="5">V</td></tr>
<tr><td>K</td></tr>
<tr><td>L</td></tr>
<tr><td>T</td></tr>
<tr><td>R</td></tr>
</table>

Since firms are price takers, industries pay the same price for the input. Thus, aggregation over industries amounts to adding up the values of the columns of U. However, since the price of commodities differ, adding up the rows is incorrect. Before aggregating over commodities, divisia price indices are developed from the underlying prices and quantity series. A divisia price index is the weighted sum of the rate of change of prices. The weights are obtained by calculating the current share of expenditure over the sample period. If w_i is the share of commodity *i* and $\frac{p_i^1 - p_i^0}{p_i^0}$ denotes the growth rate in price, then divisia price index is given by:

$$DPI = \sum_i w_i \frac{p_i^1 - p_i^0}{p_i^0}$$

Divisia index has several advantages and desirable properties. Divisia indices are chain linked, that is, for each year the current prices are used as a base to estimate the rate of growth to the following year. In addition, the index is invariant to the production configuration. Also, it is symmetric in prices and quantities and produces the same index if the data series are reversed. It is also easily computed by TSP.[64]

In aggregating the data, some simplifying assumptions were made. In all the models developed, government was not optimizing agent, nor did the model include the possibility of trade. Government spending and net imports were added to the consumption column. Non-competing imports were assumed to provide the same service as capital to the industries, omitting difference in prices, and was added to the capital rows for industries only. For consumption, divisia indices were used. These simplifications reduce the inter-industry table to the form:

[64] Time Series Processor (TSP), see Jorgenson and Griliches (1971) and Hall and Schnake (1983) for further details on divisia index.

Figure A2: Reference Inter-industry Table

U Commodity x industry 35 x 35	C	I
K	Kc	Ki
L	Lc	Li

DATA FOR CHAPTER 2

As mentioned in Chapter 2, the model is represented by a single industry. Divisia indices for price and quantity were obtained for the service sector, column 34 of U. The time series data for the price and quantity index was used for estimating the model.

DATA FOR CHAPTER 3

The model in this chapter consisted of a single producer and consumer with two input factors. The inter-industry table is aggregated to the table that takes the form:

Figure A3: Inter-industry Table for Chapter 3

U	C
K	Kc
L	Lc

Again for simplification, the investment column is added to the consumption column to form a new consumption allocation. The columns of all the industries are added to form a single industry column. Then, input divisia price and quantity indices are obtained to provide the representative industry and consumption.

The time series of the intermediate input cost to the industry (U), capital services input cost (K), total input cost (U+K+L), consumption of good (C), and the industry output represented by the divisia quantity, completes the data for the estimation in Chapter 3.

DATE FOR CHAPTER 4

The data series for this chapter consists of two industries: energy and materials; two final demand columns: consumption and investment; and four inputs: energy, materials, capital, and labor. The aggregated inter-industry table reduces to the following form:

Figure A4: Inter-industry Table for Chapter 4

EE	EM	EC	EI
ME	MM	MC	MI
KE	KM	KC	KI
LE	LM	LC	LI

Where: EE is energy input to energy industry, EM is material input to materials industry, and so on.

Five industries (coal mining, crude petroleum and natural gas, petroleum refining, electric utilities, and gas utilities) are aggregated to form the Energy Industry. The remaining 30 industries are aggregated to form the Materials Industry (see Table A1).

The aggregation steps are similar to the previous discussion. Starting with the inter-industry table given in Figure A1, the corresponding industry columns for the Energy and Materials are added, and then the commodity rows are divisia aggregated, to get the desired data set. The time series values of cells LC, MC, and KC along with its corresponding price indices form the data set for consumer model estimation. The producer model data set includes a logarithm of the all the input shares for the three industries and its corresponding divisia price indices.

Table A1: Description of the Industries

Number	Description	Energy	Materials
1	Agriculture, forestry, and fisheries		1
2	Metal mining		2
3	Coal mining	1	
4	Crude petroleum and natural gas	2	
5	Nonmetallic mineral mining		3
6	Construction		4
7	Food and kindred product		5
8	Tobacco manufactures		6
9	Textile mill products		7
10	Apparel and other textile products		8
11	Lumber and wood products		9
12	Furniture and fixtures		10
13	Paper and allied products		11
14	Printing and publishing		12
15	Chemicals and allied products		13
16	Petroleum refining	3	
17	Rubber and plastic products		14
18	Leather and leather products		15
19	Stone, clay and glass products		16
20	Primary metals		17
21	Fabricated metal products		18
22	Machinery, except electrical		19
23	Electrical machinery		20
24	Motor vehicles		21
25	Other transportation equipment		22
26	Instruments		23
27	Miscellaneous manufacturing		24
28	Transportation and warehousing		25
29	Communication		26
30	Electric utilities	4	
31	Gas utilities	5	
32	Trade		27
33	Finance, insurance and real estate		28
34	Other services		29
35	Government enterprises		30

Appendix B: Implementation of PROTO in GEMPACK

In this appendix, user inputs and the GEMPACK routines used for calculating the base case solution, testing the validity of the model, running a policy simulation, and evaluating the Jacobian matrix is discussed.

CALCULATION OF THE BASE CASE SOLUTION IN GEMAPCK

Calculation of the base case solution in GEMPACK involves following four steps:

Step 1	Prepare initial input guess file;
Step 2	Run TABLO and GEMSIM to calculate the slack values;
Step 3	Run simulation on the "slack augmented model";
Step 4	View the base case solution.

Step 1: Prepare Initial Input Guess File

The first step is to prepare a text file containing the initial guesses of the endogenous variables, values of the exogenous variables, and parameters.[65] All the endogenous variables were initialized to 1. The sample of the GEMPACK text file (*guessfile.txt)* is attached at the end of this appendix. GEMPACK routine MODHAR is used to convert the text data file to a header array file. Although data files can be read in as a text file, the choice for holding the data in

[65]The logical initial guesses for the variables would be the present values of the economy. However, for the models illustrated in this document, some random initial guesses are chosen.

GEMPACK is a header array file.[66] The user inputs for running MODHAR is given below. The output is written to a header array file (*guessfile).*

If the above procedure needs to be performed repeatedly, the keystroke information can be stored in a stored-input file (SIF) while running the MODHAR for the first time.[67]

This user input will create a text batch file (*guessfile.sti*), which stores all the inputs needed to create a header array file (*guessfile*) from a text file (*guessfile.txt*). This procedure is easy and avoids repetition of the keystrokes. The following table shows the user input to invoke the stored-input file.

User Input to MODHAR to create header file

<carriage-return>	! Use the default options
n	! Not based on old file (create a new file)
guessfile	! Name of the header file to be created
<carriage-return>	! Use the default information for the file name
at	! Add all the arrays from the text file
guessfile.txt	! Name of the text file (created prior to running MODHAR)
a	! Add all the array on the file
	[MODHAR transfers all the arrays from the text file and creates the header file. This will take a moment or two.]
ex	! Exit with save and shows the summary of the transfer
<my name>	! This portion is to create a history
PROTO model	! write a short description
**end	! end of history
y	! Yes, this history is what I want

User Input to MODHAR to create a stored-input file

[66]GEMPACK USER DOCUMENTATION, Release 5.1 Vol 1, page 3-24. see also "How to Create and Modify GEMPACK Header Array Files Using the Program MODHAR", GEMPACK Document No. 3

[67]Stored-input file is a very useful tool, this is also available in GEMSIM. See GEMPACK documentation for details.

<carriage-return>	! Use the default options
sif	! Create a stored-input file for future use
guessfile.sti	! Input the name of the stored-input file
<carriage-return>	! Use the default options
n	! Not based on all the arrays from the text file
guessfile	! Name of the header file to be created
<carriage-return>	! Use the default information for the file name
at	! Add all the arrays from the text file
guessfile.txt	! Name of the text file to be created (prior to running MODHAR)
a	! Add all the array on the file
	[MODHAR transfers all the arrays from the text file and creates the header file. This will take a moment or two.]
ex	! Exit with save and shows the summary of the transfer
<my name>	! This portion is to create a history
PROTO model	! write a short description
**end	! end of history
y	! Yes, this history is what I want

User Input to MODHAR to invoke stored-input file

sti	! Take inputs from stored-input file
guessfile.sti	! Name of the stored-input file
	[MODHAR transfers all the arrays from the text file and creates the header file. This will take a moment or two]

Step 2: Run TABLO and GEMSIM to Calculate The Slack Values

Although TABLO and GEMSIM are primarily used to perform simulation on a model, they can also be used manipulate data. In this section, TABLO is used to manipulate the data from the header array file (*guessfile*). The initial values of the endogenous variables and the values of the parameters are read in from the header file and then substituted into the "slack augmented model" to calculate the values of the slack variables. GEMPACK program file (*slackcal.tab*) is attached at the end of the appendix. The initial guesses along with the calculated slack values solve the "slack augmented model."

The user inputs shown in the next page can be avoided by simply calling the routine TABLOX. TABLOX routine linearizes the nonlinear model and generates Fortran files for GEMSIM. GEMSIM routine calls in all the specified data files into the file generated by TABLOX and calculates the slack values. The slack values are then stored internally in a header array file (*slackfile*). The user inputs for TABLO and GEMSIM are given below.

User Input to TABLO

<carriage-return>	! Use the default options
Slackcal	! Name of the TABLO input file
	[TABLO performs checks on formulas, equations. If there are no syntax errors, the next command is requested.]
a	! begin automatic code generation
wfp	! write the Fortran program
<carriage-return>	! use the other default code generation options
slackfile	! name of the program to be written
	[TABLO will generate *slackcal.f, slackcal.axt, slackcal.axs* files, TABLO- generated program.]

User Input to GEMSIM

<carriage-return>	! Use the default options
slackfile	! Name of the GEMSIM auxiliary statement file
guessfile	! Name of the input header file
slackfile	! Name of the file to which the slacks will be stored
	[GEMSIM calculates the slack values and stores in the file specified above.]

Note that the user inputs could have been placed in a stored-input file for future use. The stored-input file, *slackcal.sti*, is also attached at the end.

Step 3: Run Simulation On The Slacks

So far, only the "slack augmented model" is solved. The objective of this step is to run an experiment where the slack variables are driven down to zero. The experiment yields the solution for the original model. The TABLO file (*sim.tab*) specifies the model and the pointers to files to be accessed. Although the model equations are the same as in the program *simcal.tab*, program *sim.cal* differs in other parts of the program.

GEMPACK routine GEMSIM specifies the data files, model closure, and the types of variables.[68] All the information needed for GEMSIM program to run is written in the command file (*simslack.cmf*). The purpose of this command file is to avoid re-entering the commands manually every time GEMSIM is called.[69]

The solution to the simulation in percentage change is written in a GEMPACK solution file (*simslack.sl4*). The base case solution in levels is updated in the header file *baseline*. The TABLO file and GEMSIM commands for this step are attached at the end. The user inputs are illustrated below.

---

[68]GEMPACK USER DOCUMENTATION, Release 5.1, page 2-3

[69]The command file is similar to stored-input file. In a command file the information need not follow a prescribed sequence, whereas for stored-input file the input sequence must match the input requested by TABLO or GEMSIM.

User Input to TABLO

<carriage-return>	! Use the default options
sim	! Name of the TABLO input file
<carriage-return>	! Use the default information file name
	[TABLO performs checks on formulas, equations. If there are no syntax errors, the next command is requested.]
a	! Begin automatic code generation
wfp	! Write the Fortran program
sim	! Name of the program to be written
	[TABLO will generate *sim.f, sim.axt, sim.axs* files, TABLO- generated program.]

Again, it is possible to perform the same operation by typing TABLOX *sim*.

User Input to GEMSIM

<carriage-return>	! Use the default options
cmf	! Take input from a command file
simslack.cmf	! Name of the command input file
	[GEMSIM computes the solution and writes the result in a solution file *slacksim.sl4*. The solution is in percentage change. The updated level values are written to the file specified in the command file. In this example its *baseline*.]

As in TABLO, it is possible to invoke the command file by typing GEMSIMX *simslack.*

Step 4: View the Base Case Solution

The solution of the original model is stored in a header file specified in the command file. SEEHAR routine converts the header array file into a text file.

The base case solution is stored in a text file (*baseline.txt*). A sample of the output is attached at the end. The user inputs for this procedure is shown below.

User Input to SEEHAR

ss	! Spreadsheet output
x	! Space separator
<carriage-return>	! Use the default options
baseline	! Name of the header array file to examine
baseline.txt	! Name of the text file to convert the header file
<carriage-return>	! Accept the name of the file
r	! output all the variables or other options from the menu
	[SEEHAR writes the updated header array file to the specified text file]

GEMPACK also stores the solution in percentage change in the solution file (*simslack.sl4*). This can be viewed as a text file by running GEMPACK routine GEMPIE. The output from the GEMPIE is written to *simslack.pi5*.

MODEL VALIDITY: HOMOGENEITY TEST

Performing the homogeneity test checks the validity of the model. The model is valid if the homogeneity properties of the demand and expenditure functions are satisfied. Running a simulation, in which the price of the consumption good, the numeraire, is increased by 100 percent, simulates the homogeneity test. This experiment should increase income, wage and subsidy by 100 percent, and consumption, labor, leisure and output should remain unchanged.

A command file (*homog.cmf*) with appropriate file names and the model closure is created. It is invoked by GEMSIMX routine. The simulation results are stored in the solution file (*homog.sl4)*. GEMPIE routine converts the solution file into a text file. The user inputs are shown below.

User Input to GEMPIE

<carriage-return>	! Use the default options
homog	! Name of the solution file
a	! All components of all cumulatively retained endogenous
variable	
<carriage-return>	! Use the default options
homg.txt	! Type the heading or title
6	! Number of decimal places wanted
	[GEMPIE writes the solution file to a text file, *homog.pi5*]

The results for this homogeneity experiment are logged in the GEMPIE output file (*homog.pi5*). The results indicate that the percentage changes of the variables are approximately equal to the expected result.

Homogeneity Results

Endogenous variable	Expected Percent Change (%)	Percent Change (%)
P	-	100.00
C	-0.000007	0.00
J	0.000010	0.00
L	0.00	0.00
Pk	99.999954	100.000015
Q	-0.000037	0.00
S	99.999992	99.9999
W	100.00	100.00
X	-0.000037	-0.00003
Y	99.999992	100.00

POLICY SIMULATION

The base case solution gives the starting point for all the simulations. Given the base case solution, policy simulations in GEMPACK amount to only changing the closure of the model. Tax on consumption is increased from the present value of 10 percent to 20 percent. That is, in terms of percentage change, the tax is increased by 100 percent. The command file for GEMSIM (*simtau.cmf*), is attached at the end of this appendix. The procedure to invoke the simulation is the same as described in the previous section.

User Input to GEMSIM

<carriage-return>	! Use the default options
cmf	! Take input from the command file
simtau.cmf	! Name of the command input file
	[GEMSIM computes the solution and writes the result in a solution file, *simtau.sl4*. The solution is in percentage change. The updated level values are written to the file specified in the command file. In this example it is *simtau.upd*]

The results of the tax simulation are reported in the solution file (*simtau.sl4*). To run different simulations, one need to only change the model closure and the shock specification in the GEMSIM command file.

EVALUATION OF THE JACOBIAN MATRIX

The procedure described in the above section includes only one shocked variable. However, GEMPACK routine SAGEM enables the model to perform multiple shocks within a single run. This is important for evaluating the inverse

of a matrix in GEMPACK. All four parameters are increased by one percent simultaneously and the resulting solution file (*protosagem.sl4*) gives the inverse for the PROTO model. Routine GEMPIE is used to output the inverse into a text file. The command file (*protosagem.cmf*) for the simulation is attached at the end, and user inputs are shown below.

User Input to SAGEM

<carriage-return>	! Use the default options
cmf	! Take input from the command file
protosagem.cmf	! Name of the command input file
	[SAGEM computes the solution and writes the result in a solution file, *protosagem.sl4*. The solution file can be viewed by calling GEMPIE routine]

Confidence interval calculation in GAUSS

All the confidence interval calculations are done using GAUSS. the Jacobian matrix, covariance matrix, and base case solution are loaded from an external file and multiplied appropriately to get the solution. The GAUSS code for determining confidence intervals for the PROTO model is attached at the end of Appendix B.

```
/*************************************************************/
/*** Gauss program to determine confidence interval        ***/
/*** Jacobian matrix is determined using GEMPACK SAGEM     ***/
/*************************************************************/
output file = ciproto.out on;

n = 11;                  /*** number of endogenous variables ***/
m = 4;                   /*** number of parameters           ***/

A = zeros(n,n);          /*** parameter coef matrix          ***/
B = zeros(n,m);          /*** parameter coef matrix          ***/
bvec = zeros(m,1);       /*** vector of parameter values     ***/
avec=zeros(n,1);         /*** vector of endogenous variables ***/
bsmall = zeros(m,m);
asmall = zeros(n,n);
vcov  = zeros(m,m) ;     /*** variance matrix for the parameters */
tmpcov  = zeros(m,m) ;   /*** variance matrix for the parameters */

/* read the covariance matrix */
load vcov[m,m] = fimlvcov.txt;


/* read jacobian matrix and baseline solution and parameters */
load ainvb[11,4] = ainvb.inv;

/*** Base case solution ***/
 c       =    2.273833E+06 ;
 j       =    7.581553E+06 ;
 k       =     928003.         ;
 l       =    1.327446E+06 ;
 pk      =    0.855376      ;
 q       =    4.487664E+06 ;
 s       =     227383.         ;
 w       =     1.11495        ;
 x       =    2.213832E+06 ;
 y       =    1.095430E+07 ;
 walras =   -0.661273     ;
 h       =    8.909000E+06 ;
 p       =     1.00000     ;
 t       =    0.100000     ;
 alpha   =     0.493315     ;
```

```
beta    =    0.176883    ;
delta   =    0.228332    ;
atech   =     2.79849    ;

avec[1,1] = c;
avec[2,1] = j;
avec[3,1] = k;
avec[4,1] = l;
avec[5,1] = pk;
avec[6,1] = q;
avec[7,1] = s;
avec[8,1] = w;
avec[9,1] = x;
avec[10,1] = y;
avec[11,1] = walras;

bvec[1,1] = alpha ;
bvec[2,1] = beta;
bvec[3,1] = gamma1;
bvec[4,1] = atech;

/*** the model is A*xhat + D*dhat = B*bhat  ***/
/*** xhat is endogenous                    ***/
/*** Var(xhat) =  BBAR*Varinvb*BBAR        ***/
/*** initialize the matrix                 ***/

bhat = ainvb*inv(diagrv(bsmall,bvec));
vendo = bhat*vcov*bhat';
varx = diagrv(asmall,avec)*bhat*vcov*(diagrv(asmall,avec)*bhat)';
/*** 95 percent confidence interval in percentage ***/
sigvendo = 2*sqrt(diag(vendo))*100;
/*** 95 percent confidence interal in levels ***/
sigvarx = 2*sqrt(diag(varx));
print sigvendo;
print sigvarx;
```

Appendix C: GEMPACK Programs for The PROTO MODEL

```
!-------------------------------------------------------------------------------------------!
! guessfile.txt                                                                             !
!-------------------------------------------------------------------------------------------!
! This file contains the initial parameter values and                                       !
! the initial guesses for the variables
!-------------------------------------------------------------------------------------------!
1 header "C" longname "consumption";
2000000
1 header "J" longname "leisure";
7000000
1 header "K" longname "capital";
928000
1 header "L" longname "labor";
1000000
1 header "Pk" longname "rental price of capital";
1.0
1 header "Q" longname "output";
5000000
1 header "S" longname "subsidy";
260000
1 header "W" longname "wage";
1
1 header "X" longname "intermediate good";
2000000
1 header "Y" longname "full income";
10000000
1 header "wal" longname "walras";
1
1 header "H" longname "time endowment";
8909000
1 header "P" longname "price of output";
1
1 header "t" longname "tax on consumption";
0.1
1 header "apha" longname "alpha";
0.493315
1 header "beta" longname "beta";
0.176883
```

1 header "gama" longname "gamma";
0.228332
1 header "atec" longname "atech";
2.79849

```
!-------------------------------------------------------------------------------------------!
! slackcal.tab                                                                              !
! This purpose of this file is calculate the slack values and                               !
! write it to a header array file (slackfile)                                               !
!-------------------------------------------------------------------------------------------!
! external file definition                                                                  !
!-------------------------------------------------------------------------------------------!
file iodata;
file (new) slackfile;
!-------------------------------------------------------------------------------------------!
! Coefficient Definition                                                                    !
!-------------------------------------------------------------------------------------------!
coefficient  c ;
coefficient  j ;
coefficient  k ;
coefficient  l ;
coefficient  pk;
coefficient  q ;
coefficient  s ;
coefficient  w ;
coefficient  x ;
coefficient  y ;
coefficient  walras;
coefficient  h ;
coefficient  p ;
coefficient  t ;
coefficient  alpha ;
coefficient  beta  ;
coefficient  gamma ;
coefficient  atech ;
!-------------------------------------------------------------------------------------------!
! Slack Definition                                                                          !
!-------------------------------------------------------------------------------------------!
coefficient  slack1 ;
coefficient  slack2 ;
coefficient  slack3 ;
coefficient  slack4 ;
coefficient  slack5 ;
coefficient  slack6 ;
coefficient  slack7 ;
coefficient  slack8 ;
```

```
coefficient  slack9 ;
coefficient  slack10 ;
coefficient  slack11 ;
!-------------------------------------------------------------------------------------------!
! Read initila values from external fies                                                    !
!-------------------------------------------------------------------------------------------!
read c from file iodata header "C" ;
read j from file iodata header "J" ;
read k from file iodata header "K" ;
read l from file iodata header "L" ;
read pk from file iodata header "Pk" ;
read q from file iodata header "Q" ;
read s from file iodata header "S" ;
read w from file iodata header "W" ;
read x from file iodata header "X" ;
read y from file iodata header "Y" ;
read walras from file iodata header "wal" ;
read h from file iodata header "H" ;
read p from file iodata header "P" ;
read t from file iodata header "t" ;
read alpha from file iodata header "apha" ;
read beta from file iodata header "beta" ;
read gamma from file iodata header "gama" ;
read atech from file iodata header "atec" ;
!-------------------------------------------------------------------------------------------!
! Slack Augmented Model                                                                     !
!-------------------------------------------------------------------------------------------!
formula   slack1 = - p*c*(1+t) + gamma*y ;
formula   slack2 = -w*j + (1-gamma)*y ;
formula   slack3 = -y + w*h + pk*k + s ;
formula   slack4 = -h + l + j ;
formula   slack5 =  - x + alpha*q ;
formula  slack6 = - k*pk + beta*p*q ;
formula  slack7 = - l*w + (1-alpha-beta)*p*q ;
formula  slack8 = - q + atech*(x^alpha)*(k^beta)*(l^(1-alpha-beta)) ;
formula  slack9 = - t*p*c + s   ;
formula  slack10 = - walras + q - x - c  ;
formula  slack11 = - k + 928003   ;
!-------------------------------------------------------------------------------------------!
! Calculate the values of the slacks and dump it out                                        !
!-------------------------------------------------------------------------------------------!
write slack1 to file slackfile header "s001" longname "slack1" ;
```

```
write slack2 to file slackfile header "s002" longname "slack2" ;
write slack3 to file slackfile header "s003" longname "slack3" ;
write slack4 to file slackfile header "s004" longname "slack4" ;
write slack5 to file slackfile header "s005" longname "slack5" ;
write slack6 to file slackfile header "s006" longname "slack6" ;
write slack7 to file slackfile header "s007" longname "slack7" ;
write slack8 to file slackfile header "s008" longname "slack8" ;
write slack9 to file slackfile header "s009" longname "slack9" ;
write slack10 to file slackfile header "s010" longname "slack10" ;
write slack11 to file slackfile header "s011" longname "slack11" ;
```

```
!-------------------------------------------------------------------------------------------!
! simslack.tab                                                                              !
! Tablo file of the PROTO model                                                             !
!-------------------------------------------------------------------------------------------!
equation          ( default = levels )   ;
variable          ( default = levels )   ;
coefficient       ( default = parameter )   ;
formula        ( default = initial )    ;
!-------------------------------------------------------------------------------------------!
! Variable Definition                                                                       !
!-------------------------------------------------------------------------------------------!
variable  c ;
variable  j ;
variable  k ;
variable  l ;
variable  pk;
variable  q ;
variable  s ;
variable  w ;
variable  x ;
variable  y ;
variable  walras;
variable  h ;
variable  p ;
variable  t ;
variable  alpha ;
variable  beta  ;
variable  gamma ;
variable  atech ;
variable  slack1 ;
variable  slack2 ;
variable  slack3 ;
variable  slack4 ;
variable  slack5 ;
variable  slack6 ;
variable  slack7 ;
variable  slack8 ;
variable  slack9 ;
variable  slack10 ;
variable  slack11 ;
!-------------------------------------------------------------------------------------------!
! External file definition                                                                  !
```

```
!-------------------------------------------------------------------------------------------!
file iodata ;
file slackfile ;
!-------------------------------------------------------------------------------------------!
! read in the initial guesses for the variables                                             !
!-------------------------------------------------------------------------------------------!
read c from file iodata header "C" ;
read j from file iodata header "J" ;
read k from file iodata header "K" ;
read l from file iodata header "L" ;
read pk from file iodata header "Pk" ;
read q from file iodata header "Q" ;
read s from file iodata header "S" ;
read w from file iodata header "W" ;
read x from file iodata header "X" ;
read y from file iodata header "Y" ;
read walras from file iodata header "wal" ;
read h from file iodata header "H" ;
read p from file iodata header "P" ;
read t from file iodata header "t" ;
read alpha from file iodata header "apha" ;
read beta from file iodata header "beta" ;
read gamma from file iodata header "gama" ;
read atech from file iodata header "atec" ;
!-------------------------------------------------------------------------------------------!
! Read in the slack values                                                                  !
!-------------------------------------------------------------------------------------------!
read slack1 from file slackfile header "s001" ;
read slack2 from file slackfile header "s002" ;
read slack3 from file slackfile header "s003" ;
read slack4 from file slackfile header "s004" ;
read slack5 from file slackfile header "s005" ;
read slack6 from file slackfile header "s006" ;
read slack7 from file slackfile header "s007" ;
read slack8 from file slackfile header "s008" ;
read slack9 from file slackfile header "s009" ;
read slack10 from file slackfile header "s010" ;
read slack11 from file slackfile header "s011" ;
!-------------------------------------------------------------------------------------------!
! Model Equations                                                                           !
!-------------------------------------------------------------------------------------------!
equation eq1   slack1 = - p*c*(1+t) + gamma*y ;
```

```
equation eq2   slack2 = -w*j + (1-gamma)*y ;
equation eq3   slack3 = -y + w*h + pk*k + s ;
equation eq4   slack4 = -h + l + j ;
equation eq5   slack5 =  - x + alpha*q ;
equation eq6  slack6 = - k*pk + beta*p*q ;
equation eq7  slack7 = - l*w + (1-alpha-beta)*p*q ;
equation eq8  slack8 = - q + atech*(x^alpha)*(k^beta)*(l^(1-alpha-beta)) ;
equation eq9  slack9 = - t*p*c + s   ;
equation eq10  slack10 = - walras + q - x - c  ;
equation eq11  slack11 = - k + 928003   ;
```

```
!-------------------------------------------------------------------------------------------!
! simslack.cmf                                                                              !
! Input for GEMSIM for simulating slack experiment                                          !
!-------------------------------------------------------------------------------------------!
auxiliary file = simslack;
file iodata = guessfile;
file slackfile = slackfile;
updated file iodata = basecase;
updated file slackfile = slackfile.upd;
intermediate extra data = simslack ;
extrapolation accuracy file = yes;
method = euler;
steps  = 3;
model = simslack ;
version = 1;
identifier = stylized johansen, standard data;
verbal description = euler model ;
equation file = simslack;
solution file = simslack;
log only = simslack.log;

endogenous p_c;
endogenous p_j;
endogenous p_k;
endogenous p_l;
endogenous p_pk;
endogenous p_q;
endogenous p_s;
endogenous p_w;
endogenous p_x;
endogenous p_y;
endogenous p_walras;

rest exogenous ;

shock p_slack1 = uniform -100 ;
shock p_slack2 = uniform -100 ;
shock p_slack3 = uniform -100 ;
shock p_slack4 = uniform -100 ;
shock p_slack5 = uniform -100 ;
shock p_slack6 = uniform -100 ;
shock p_slack7 = uniform -100 ;
```

```
shock p_slack8 = uniform -100 ;
shock p_slack9 = uniform -100 ;
shock p_slack10 = uniform -100 ;
shock p_slack11 = uniform -100 ;

save environment file protoenv;
cpu =yes;
```

```
!-------------------------------------------------------------------------------------------!
! simtau.cmf                                                                                !
! Input for GEMSIM for simulating tax experiment                                            !
!-------------------------------------------------------------------------------------------!
auxiliary file = simslack;
file iodata = basecase;
file slackfile = slackfile;
updated file iodata = simtau;
updated file slackfile = slackfile.upd;
intermediate extra data = simtau ;
extrapolation accuracy file = yes;
method = euler;
steps  = 1 ;
model = simtau ;
version = 1;
identifier = stylized johansen, standard data;
verbal description = euler model ;
equation file = simtau;
solution file = simtau;
log only = simtau.log;

endogenous p_c;
endogenous p_j;
endogenous p_k;
endogenous p_l;
endogenous p_pk;
endogenous p_q;
endogenous p_s;
endogenous p_w;
endogenous p_x;
endogenous p_y;
endogenous p_walras;

rest exogenous ;

shock p_t = uniform 100 ;

cpu =yes;
```

```
!-------------------------------------------------------------------------------------------!
! protosagem.cmf                                                                            !
! Input for SAGEM for simulating an experiment to get the Jacobian Matrix       !
!-------------------------------------------------------------------------------------------!
use equation file proto;
solution file = protosagem;
use environment file protoenv;

shock p_alpha = uniform 1;
shock p_beta = uniform 1;
shock p_gamma = uniform 1;
shock p_atech = uniform 1;

verbal description = To get inv(A)*B ;

individually-retained exogenous %all;
cpu =yes;
```

```
!-------------------------------------------------------------------------------------------!
! homog.cmf                                                                                 !
! Input for GEMSIM for simulating homgeneity test                                           !
!-------------------------------------------------------------------------------------------!
auxiliary file = simslack;
file iodata = basecase;
file slackfile = slackfile;
updated file iodata = homog;
updated file slackfile = slackfile.upd;
intermediate extra data = homog ;
extrapolation accuracy file = yes;
method = euler;
steps  = 1 ;
model = homog ;
version = 1;
identifier = stylized johansen, standard data;
verbal description = euler model ;
equation file = homog;
solution file = homog;
log only = homog.log;

endogenous p_c;
endogenous p_j;
endogenous p_k;
endogenous p_l;
endogenous p_pk;
endogenous p_q;
endogenous p_s;
endogenous p_w;
endogenous p_x;
endogenous p_y;
endogenous p_walras;
rest exogenous ;

shock p_p = uniform 100 ;
cpu =yes;
```

Appendix D: GEMPACK TABLO File For The Intertemporal Model

```
!-------------------------------------------------------------------------------------------!
! Intertemporal Model                                                                       !
!-------------------------------------------------------------------------------------------!
equation        ( default = levels );
variable        ( default = levels );
coefficient     ( default = parameter );
formula         ( default = initial );
!-------------------------------------------------------------------------------------------!
! Time Horizon                                                                              !
!-------------------------------------------------------------------------------------------!
set (intertemporal) alltime      maximum size 30 ( p[0] - p[29] );
set (intertemporal) fwdtime    maximum size 29 ( p[0] - p[28] );
set (intertemporal) backtime   maximum size 29 ( p[1] - p[29] );
set (intertemporal) begintime  maximum size 1   ( p[0] );
set (intertemporal) endtime    maximum size  1   ( p[29] );

subset fwdtime   is subset of alltime;
subset backtime  is subset of alltime;
subset begintime is subset of alltime;
subset endtime    is subset of alltime;
!-------------------------------------------------------------------------------------------!
! External File Handles                                                                     !
!-------------------------------------------------------------------------------------------!
file iodataparm ;
file (text) time ;
file basedata ;
file slackfile ;
!-------------------------------------------------------------------------------------------!
! Variable Definition                                                                       !
!-------------------------------------------------------------------------------------------!
variable (all,i,alltime) share_KE(i)  ;
variable (all,i,alltime) share_LE(i)  ;
variable (all,i,alltime) share_EE(i)  ;
variable (all,i,alltime) share_ME(i)  ;
variable (all,i,alltime) share_KM(i)  ;
variable (all,i,alltime) share_LM(i)  ;
variable (all,i,alltime) share_EM(i)  ;
```

```
variable (all,i,alltime) share_MM(i)  ;
variable (all,i,alltime) share_KI(i)  ;
variable (all,i,alltime) share_LI(i)  ;
variable (all,i,alltime) share_EI(i)  ;
variable (all,i,alltime) share_MI(i)  ;
variable (all,i,alltime) pIindex(i)  ;
variable (all,i,alltime) pKindex(i)  ;
variable (all,i,alltime) pEindex(i)  ;
variable (all,i,alltime) pMindex(i)  ;
variable (all,i,alltime) phatEindex(i)  ;
variable (all,i,alltime) phatMindex(i)  ;
variable (all,i,alltime) income(i)  ;
variable (all,i,alltime) sup_income(i)  ;
variable (all,i,alltime) sup_expend(i)  ;
variable (all,i,alltime) int_rate(i)  ;
variable (all,i,alltime) qKindex(i)  ;
variable (all,i,alltime) qEindex(i)  ;
variable (all,i,alltime) qMindex(i)  ;
variable (all,i,alltime) qGindex(i)  ;
variable (all,i,alltime) qIindex(i)  ;
variable (all,i,alltime) qKEindex(i)  ;
variable (all,i,alltime) qLEindex(i)  ;
variable (all,i,alltime) qEEindex(i)  ;
variable (all,i,alltime) qMEindex(i)  ;
variable (all,i,alltime) qMMindex(i)  ;
variable (all,i,alltime) qKMindex(i)  ;
variable (all,i,alltime) qLMindex(i)  ;
variable (all,i,alltime) qEMindex(i)  ;
variable (all,i,alltime) qKCindex(i)  ;
variable (all,i,alltime) qLCindex(i)  ;
variable (all,i,alltime) qECindex(i)  ;
variable (all,i,alltime) qMCindex(i)  ;
variable (all,i,alltime) qKIindex(i)  ;
variable (all,i,alltime) qLIindex(i)  ;
variable (all,i,alltime) qEIindex(i)  ;
variable (all,i,alltime) qMIindex(i)  ;
variable (all,i,alltime) walras(i)  ;
variable (all,i,alltime) subsidy(i)  ;
variable (all,i,alltime) pvwealth(i)  ;
variable (all,i,alltime) expend(i)  ;
variable (all,i,alltime) savings(i)  ;
variable (all,i,alltime) pLindex(i)  ;
```

```
variable (all,i,alltime) qLindex(i)  ;
variable (all,i,alltime) tax_K(i)  ;
variable (all,i,alltime) tax_L(i)  ;
variable (all,i,alltime) tax_E(i)  ;
variable (all,i,alltime) tax_M(i)  ;
variable (all,i,alltime) tax_income(i)  ;
variable (all,i,alltime) sv(i)  ;
variable (all,i,alltime) omega(i)  ;
variable (all,i,alltime) pvwhat(i)  ;
variable (all,i,alltime) lw(i)  ;
variable (all,i,alltime) phat(i)  ;
variable (all,i,alltime) V(i)  ;
variable (all,i,alltime) theta(i)  ;
variable (all,i,alltime) phi(i)  ;
variable (all,i,alltime) utility(i)  ;
variable  qKindex0 ;
!-------------------------------------------------------------------------------------------!
! Coefficient Definition                                                                    !
!-------------------------------------------------------------------------------------------!
variable gamma_KE       ;
variable gamma_LE       ;
variable gamma_EE       ;
variable SIGMA_E ;
variable AE     ;
variable gamma_KM       ;
variable gamma_LM       ;
variable gamma_EM       ;
variable SIGMA_M ;
variable AM     ;
variable gamma_KI       ;
variable gamma_LI       ;
variable gamma_EI       ;
variable SIGMA_I ;
variable AI     ;
variable cbare   ;
variable cbarm   ;
variable cbark   ;
variable cbarl   ;
variable alphal  ;
variable alphae  ;
variable alpham  ;
variable alphak  ;
```

```
!-----------------------------------------------------------------------------------------!
! Error Variables for the Econometric Model Equations                                     !
!-----------------------------------------------------------------------------------------!
variable err1 ;
variable err2 ;
variable err3 ;
variable err4 ;
variable err5 ;
variable err6 ;
variable err7 ;
variable err8 ;
variable err9 ;
variable err10 ;
variable err11 ;
variable err12 ;
variable err13 ;
variable err14 ;
variable err15 ;
variable err16 ;
variable err17 ;
variable err18 ;
!-----------------------------------------------------------------------------------------!
! Exogenous Variable Definition                                                           !
!-----------------------------------------------------------------------------------------!
coefficient adjust_pk;
coefficient dep_rate ;
coefficient disc_rate;
coefficient (all,i,fwdtime) dt(i);
coefficient (all,i,alltime) year(i);
!-----------------------------------------------------------------------------------------!
! Slack Variable Definition                                                               !
!-----------------------------------------------------------------------------------------!
variable (all,t,alltime) slack1(t) ;
variable (all,t,alltime) slack2(t) ;
variable (all,t,alltime) slack3(t) ;
variable (all,t,alltime) slack4(t) ;
variable (all,t,alltime) slack5(t) ;
variable (all,t,alltime) slack6(t) ;
variable (all,t,alltime) slack7(t) ;
variable (all,t,alltime) slack8(t) ;
variable (all,t,alltime) slack9(t) ;
variable (all,t,alltime) slack10(t) ;
```

```
variable (all,t,alltime) slack11(t) ;
variable (all,t,alltime) slack12(t) ;
variable (all,t,alltime) slack13(t) ;
variable (all,t,alltime) slack14(t) ;
variable (all,t,alltime) slack15(t) ;
variable (all,t,alltime) slack16(t) ;
variable (all,t,alltime) slack17(t) ;
variable (all,t,alltime) slack18(t) ;
variable (all,t,alltime) slack19(t) ;
variable (all,t,alltime) slack20(t) ;
variable (all,t,alltime) slack21(t) ;
variable (all,t,alltime) slack22(t) ;
variable (all,t,alltime) slack23(t) ;
variable (all,t,alltime) slack24(t) ;
variable (all,t,alltime) slack25(t) ;
variable (all,t,alltime) slack26(t) ;
variable (all,t,alltime) slack27(t) ;
variable (all,t,alltime) slack28(t) ;
variable (all,t,alltime) slack29(t) ;
variable (all,t,alltime) slack30(t) ;
variable (all,t,alltime) slack31(t) ;
variable (all,t,alltime) slack32(t) ;
variable (all,t,alltime) slack33(t) ;
variable (all,t,alltime) slack34(t) ;
variable (all,t,alltime) slack35(t) ;
variable (all,t,alltime) slack36(t) ;
variable (all,t,alltime) slack37(t) ;
variable (all,t,alltime) slack38(t) ;
variable (all,t,alltime) slack39(t) ;
variable (all,t,alltime) slack40(t) ;
variable (all,t,alltime) slack41(t) ;
variable (all,t,alltime) slack42(t) ;
variable (all,t,alltime) slack43(t) ;
variable (all,t,alltime) slack44(t) ;
variable (all,t,alltime) slack45(t) ;
variable (all,t,alltime) slack46(t) ;
variable (all,t,fwdtime) slack47(t) ;
variable (all,t,fwdtime) slack48(t) ;
variable (all,t,begintime) slack49(t) ;
variable (all,t,endtime) slack50(t) ;
variable (all,t,alltime) slack51(t) ;
variable (all,t,fwdtime) slack52(t) ;
```

```
variable (all,t,endtime) slack53(t) ;
variable (all,t,alltime) slack54(t) ;
variable (all,t,alltime) slack55(t) ;
variable (all,t,alltime) slack56(t) ;
variable (all,t,fwdtime) slack57(t) ;
variable (all,t,endtime) slack58(t) ;
variable (all,t,fwdtime) slack59(t) ;
variable (all,t,endtime) slack60(t) ;
variable (all,t,fwdtime) slack61(t) ;
variable (all,t,endtime) slack62(t) ;
variable (all,t,alltime) slack63(t) ;
variable (all,t,alltime) slack64(t) ;
variable (all,t,alltime) slack65(t) ;
variable (all,t,alltime) slack66(t) ;
variable (all,t,alltime) slack67(t) ;
variable (all,t,alltime) slack68(t) ;
variable (all,t,alltime) slack69(t) ;
variable (all,t,alltime) slack70(t) ;
variable (all,t,alltime) slack71(t) ;
variable (all,t,alltime) slack72(t) ;
variable (all,t,alltime) slack73(t) ;
variable (all,t,alltime) slack74(t) ;
variable (all,t,alltime) slack75(t) ;
variable (all,t,alltime) slack76(t) ;
variable (all,t,alltime) slack77(t) ;
variable (all,t,alltime) slack78(t) ;
variable (all,t,alltime) slack79(t) ;
variable (all,t,alltime) slack80(t) ;
variable (all,t,alltime) slack81(t) ;
variable (all,t,alltime) slack82(t) ;
variable (all,t,alltime) slack83(t) ;
variable (all,t,alltime) slack84(t) ;
variable (all,t,alltime) slack85(t) ;
variable (all,t,alltime) slack86(t) ;
variable (all,t,alltime) slack87(t) ;
variable (all,t,alltime) slack88(t) ;
variable (all,t,alltime) slack89(t) ;
variable (all,t,alltime) slack90(t) ;
variable (all,t,alltime) slack91(t) ;
variable (all,t,alltime) slack92(t) ;
variable (all,t,alltime) slack93(t) ;
variable (all,t,alltime) slack94(t) ;
```

```
variable (all,t,alltime) slack95(t) ;
variable (all,t,alltime) slack96(t) ;
variable (all,t,alltime) slack97(t) ;
variable (all,t,alltime) slack98(t) ;
variable (all,t,alltime) slack99(t) ;
variable (all,t,alltime) slack100(t) ;
variable (all,t,alltime) slack101(t) ;
variable (all,t,alltime) slack102(t) ;
variable (all,t,alltime) slack103(t) ;
variable (all,t,alltime) slack104(t) ;
variable (all,t,alltime) slack105(t) ;
variable (all,t,alltime) slack106(t) ;
variable (all,t,alltime) slack107(t) ;
variable (all,t,alltime) slack108(t) ;
variable (all,t,alltime) slack109(t) ;
variable (all,t,alltime) slack110(t) ;
variable (all,t,alltime) slack111(t) ;
variable (all,t,alltime) slack112(t) ;
variable (all,t,alltime) slack113(t) ;
variable (all,t,fwdtime) slack114(t) ;
variable (all,t,fwdtime) slack115(t) ;
variable (all,t,begintime) slack116(t) ;
variable (all,t,endtime) slack117(t) ;
variable (all,t,alltime) slack118(t) ;
variable (all,t,fwdtime) slack119(t) ;
variable (all,t,endtime) slack120(t) ;
variable (all,t,alltime) slack121(t) ;
variable (all,t,alltime) slack122(t) ;
variable (all,t,alltime) slack123(t) ;
variable (all,t,fwdtime) slack124(t) ;
variable (all,t,endtime) slack125(t) ;
variable (all,t,alltime) slack126(t) ;
variable (all,t,alltime) slack127(t) ;
variable (all,t,alltime) slack128(t) ;
variable (all,t,alltime) slack129(t) ;
variable (all,t,alltime) slack130(t) ;
variable (all,t,alltime) slack131(t) ;
variable (all,t,alltime) slack132(t) ;
variable (all,t,alltime) slack133(t) ;
variable (all,t,alltime) slack134(t) ;
!-----------------------------------------------------------------------------------------!
! Grid Spacing                                                                            !
```

```
!-------------------------------------------------------------------------------------------!
formula (all,t,fwdtime) dt(t) = year(t+1) - year(t) ;
!-------------------------------------------------------------------------------------------!
! Read Exogenous Variables                                                                  !
!-------------------------------------------------------------------------------------------!
read adjust_pk          from file iodataparm header "E001" ;
read dep_rate           from file iodataparm header "E002" ;
read disc_rate          from file iodataparm header "E003" ;
read year               from file time ;
!-------------------------------------------------------------------------------------------!
! Read Endogenous Variables                                                                 !
!-------------------------------------------------------------------------------------------!
read share_KE           from file basedata header "v001" ;
read share_LE           from file basedata header "v002" ;
read share_EE           from file basedata header "v003" ;
read share_ME           from file basedata header "v004" ;
read share_KM           from file basedata header "v005" ;
read share_LM           from file basedata header "v006" ;
read share_EM           from file basedata header "v007" ;
read share_MM           from file basedata header "v008" ;
read share_KI           from file basedata header "v009" ;
read share_LI           from file basedata header "v010" ;
read share_EI           from file basedata header "v011" ;
read share_MI           from file basedata header "v012" ;
read pIindex            from file basedata header "v013" ;
read pKindex            from file basedata header "v014" ;
read pEindex            from file basedata header "v015" ;
read pMindex            from file basedata header "v016" ;
read phatEindex         from file basedata header "v017" ;
read phatMindex         from file basedata header "v018" ;
read income             from file basedata header "v019" ;
read sup_income         from file basedata header "v020" ;
read sup_expend         from file basedata header "v021" ;
read int_rate           from file basedata header "v022" ;
read qKindex            from file basedata header "v023" ;
read qEindex            from file basedata header "v024" ;
read qMindex            from file basedata header "v025" ;
read qGindex            from file basedata header "v026" ;
read qIindex            from file basedata header "v027" ;
read qKEindex           from file basedata header "v028" ;
read qLEindex           from file basedata header "v029" ;
read qEEindex           from file basedata header "v030" ;
```

```
read qMEindex        from file basedata header "v031" ;
read qMMindex        from file basedata header "v032" ;
read qKMindex        from file basedata header "v033" ;
read qLMindex        from file basedata header "v034" ;
read qEMindex        from file basedata header "v035" ;
read qKCindex        from file basedata header "v036" ;
read qLCindex        from file basedata header "v037" ;
read qECindex        from file basedata header "v038" ;
read qMCindex        from file basedata header "v039" ;
read qKIindex        from file basedata header "v040" ;
read qLIindex        from file basedata header "v041" ;
read qEIindex        from file basedata header "v042" ;
read qMIindex        from file basedata header "v043" ;
read walras          from file basedata header "v044" ;
read subsidy         from file basedata header "v045" ;
read pvwealth        from file basedata header "v046" ;
read expend          from file basedata header "v047" ;
read savings         from file basedata header "v048" ;
read pLindex         from file basedata header "v049" ;
read qLindex         from file basedata header "v050" ;
read tax_K           from file basedata header "v051" ;
read tax_L           from file basedata header "v052" ;
read tax_E           from file basedata header "v053" ;
read tax_M           from file basedata header "v054" ;
read tax_income      from file basedata header "v055" ;
read sv              from file basedata header "v056" ;
read omega           from file basedata header "v057" ;
read pvwhat          from file basedata header "v058" ;
read lw              from file basedata header "v059" ;
read phat            from file basedata header "v060" ;
read V               from file basedata header "v061" ;
read theta           from file basedata header "v062" ;
read phi             from file basedata header "v063" ;
read utility         from file basedata header "v064" ;
read qKindex0        from file basedata header "v065" ;
!-----------------------------------------------------------------------------------------!
! Read Parameter Values                                                                   !
!-----------------------------------------------------------------------------------------!
read gamma_KE        from file basedata header "v066" ;
read gamma_LE        from file basedata header "v067" ;
read gamma_EE        from file basedata header "v068" ;
read SIGMA_E         from file basedata header "v069" ;
```

```
read AE                 from file basedata header "v070" ;
read gamma_KM           from file basedata header "v071" ;
read gamma_LM           from file basedata header "v072" ;
read gamma_EM           from file basedata header "v073" ;
read SIGMA_M            from file basedata header "v074" ;
read AM                 from file basedata header "v075" ;
read gamma_KI           from file basedata header "v076" ;
read gamma_LI           from file basedata header "v077" ;
read gamma_EI           from file basedata header "v078" ;
read SIGMA_I            from file basedata header "v079" ;
read AI                 from file basedata header "v080" ;
read cbare              from file basedata header "v081" ;
read cbarm              from file basedata header "v082" ;
read cbark              from file basedata header "v083" ;
read cbarl              from file basedata header "v084" ;
read alphal             from file basedata header "v085" ;
read alphae             from file basedata header "v086" ;
read alpham             from file basedata header "v087" ;
read alphak             from file basedata header "v088" ;
!-------------------------------------------------------------------------------------------!
! Read Initialized Values of the Error Variables                                            !
!-------------------------------------------------------------------------------------------!
read err1               from file basedata header "v089" ;
read err2               from file basedata header "v090" ;
read err3               from file basedata header "v091" ;
read err4               from file basedata header "v092" ;
read err5               from file basedata header "v093" ;
read err6               from file basedata header "v094" ;
read err7               from file basedata header "v095" ;
read err8               from file basedata header "v096" ;
read err9               from file basedata header "v097" ;
read err10              from file basedata header "v098" ;
read err11              from file basedata header "v099" ;
read err12              from file basedata header "v100" ;
read err13              from file basedata header "v101" ;
read err14              from file basedata header "v102" ;
read err15              from file basedata header "v103" ;
read err16              from file basedata header "v104" ;
read err17              from file basedata header "v105" ;
read err18              from file basedata header "v106" ;
!-------------------------------------------------------------------------------------------!
! Read Slack Values Values                                                                  !
```

```
!-------------------------------------------------------------------------------------------!
read slack1          from file slackfile header "s001" ;
read slack2          from file slackfile header "s002" ;
read slack3          from file slackfile header "s003" ;
read slack4          from file slackfile header "s004" ;
read slack5          from file slackfile header "s005" ;
read slack6          from file slackfile header "s006" ;
read slack7          from file slackfile header "s007" ;
read slack8          from file slackfile header "s008" ;
read slack9          from file slackfile header "s009" ;
read slack10         from file slackfile header "s010" ;
read slack11         from file slackfile header "s011" ;
read slack12         from file slackfile header "s012" ;
read slack13         from file slackfile header "s013" ;
read slack14         from file slackfile header "s014" ;
read slack15         from file slackfile header "s015" ;
read slack16         from file slackfile header "s016" ;
read slack17         from file slackfile header "s017" ;
read slack18         from file slackfile header "s018" ;
read slack19         from file slackfile header "s019" ;
read slack20         from file slackfile header "s020" ;
read slack21         from file slackfile header "s021" ;
read slack22         from file slackfile header "s022" ;
read slack23         from file slackfile header "s023" ;
read slack24         from file slackfile header "s024" ;
read slack25         from file slackfile header "s025" ;
read slack26         from file slackfile header "s026" ;
read slack27         from file slackfile header "s027" ;
read slack28         from file slackfile header "s028" ;
read slack29         from file slackfile header "s029" ;
read slack30         from file slackfile header "s030" ;
read slack31         from file slackfile header "s031" ;
read slack32         from file slackfile header "s032" ;
read slack33         from file slackfile header "s033" ;
read slack34         from file slackfile header "s034" ;
read slack35         from file slackfile header "s035" ;
read slack36         from file slackfile header "s036" ;
read slack37         from file slackfile header "s037" ;
read slack38         from file slackfile header "s038" ;
read slack39         from file slackfile header "s039" ;
read slack40         from file slackfile header "s040" ;
read slack41         from file slackfile header "s041" ;
```

```
read slack42            from file slackfile header "s042" ;
read slack43            from file slackfile header "s043" ;
read slack4 4           from file slackfile header "s044" ;
read slack45            from file slackfile header "s045" ;
read slack46            from file slackfile header "s046" ;
read slack47            from file slackfile header "s047" ;
read slack48            from file slackfile header "s048" ;
read slack49            from file slackfile header "s049" ;
read slack50            from file slackfile header "s050" ;
read slack51            from file slackfile header "s051" ;
read slack52            from file slackfile header "s052" ;
read slack53            from file slackfile header "s053" ;
read slack54            from file slackfile header "s054" ;
read slack55            from file slackfile header "s055" ;
read slack56            from file slackfile header "s056" ;
read slack57            from file slackfile header "s057" ;
read slack58            from file slackfile header "s058" ;
read slack59            from file slackfile header "s059" ;
read slack60            from file slackfile header "s060" ;
read slack61            from file slackfile header "s061" ;
read slack62            from file slackfile header "s062" ;
read slack63            from file slackfile header "s063" ;
!-------------------------------------------------------------------------------------------!
! Intertemporal Model                                                                       !
!-------------------------------------------------------------------------------------------!
! Producer Model: Energy Node                                                               !
!-------------------------------------------------------------------------------------------!

! eq. 4.50                                                                                  !
equation share_KE_eq  (all,t,alltime)
slack1(t) = - loge(share_KE(t))
            + loge(gamma_KE) + (sigma_E-1)*loge(AE)
            + (sigma_E-1)*loge(pEindex(t)/pKindex(t))
            + err1 - 1 ;

! eq. 4.51                                                                                  !
equation share_LE_eq  (all,t,alltime)
slack2(t) = - loge(share_LE(t))
            + loge(gamma_LE) + (sigma_E-1)*loge(AE)
            + (sigma_E-1)*loge(pEindex(t)/pLindex(t))
            + err2 - 1 ;
```

```
! eq. 4.52                                                                                    !
equation share_EE_eq  (all,t,alltime)
slack3(t) = - loge(share_EE(t))
              + loge(gamma_EE) + (sigma_E-1)*loge(AE)
              + (sigma_E-1)*loge(pEindex(t)/pEindex(t))
              + err3 - 1  ;

! eq. 4.53                                                                                    !
equation share_ME_eq  (all,t,alltime)
slack4(t) = - loge(share_ME(t))
              + loge(1-gamma_KE-gamma_LE-gamma_EE)
              + (sigma_E-1)*loge(AE)
              + (sigma_E-1)*loge(pEindex(t)/phatMindex(t))
              + err4 - 1 ;

!-----------------------------------------------------------------------------------------------!
! Producer Model: Materials Node                                                                !
!-----------------------------------------------------------------------------------------------!

! eq. 4.50                                                                                    !
equation share_KM_eq  (all,t,alltime)
slack5(t) = - loge(share_KM(t))
              + loge(gamma_KM) + (sigma_M-1)*loge(AM)
              + (sigma_M-1)*loge(pMindex(t)/pKindex(t))
              + err5 - 1 ;

! eq. 4.51                                                                                    !
equation share_LM_eq  (all,t,alltime)
slack6(t) = - loge(share_LM(t))
              + loge(gamma_LM) + (sigma_M-1)*loge(AM)
              + (sigma_M-1)*loge(pMindex(t)/pLindex(t))
              + err6 - 1 ;

! eq. 4.52                                                                                    !
equation share_EM_eq  (all,t,alltime)
slack7(t) = - loge(share_EM(t))
               + loge(gamma_EM) + (sigma_M-1)*loge(AM)
               + (sigma_M-1)*loge(pMindex(t)/phatEindex(t))
               + err7 - 1 ;

! eq. 4.53                                                                                    !
equation share_MM_eq  (all,t,alltime)
```

```
slack8(t) = - loge(share_MM(t))
            + loge(1-gamma_KM-gamma_LM-gamma_EM)
            + (sigma_M-1)*loge(AM)
            + (sigma_E-1)*loge(pMindex(t)/pMindex(t))
            + err8 - 1 ;
!-------------------------------------------------------------------------------------------!
! Producer Model: Capital Goods Node                                                        !
!-------------------------------------------------------------------------------------------!

! eq. 4.50                                                                                  !
equation share_KI_eq  (all,t,alltime)
slack9(t) = - loge(share_KI(t))
            + loge(gamma_KI) + (sigma_I-1)*loge(AI)
            + (sigma_I-1)*loge(pIindex(t)/pKindex(t))
            + err9 - 1 ;

! eq. 4.51                                                                                  !
equation share_LI_eq  (all,t,alltime)
slack10(t) = - loge(share_LI(t))
             + loge(gamma_LI) + (sigma_I-1)*loge(AI)
             + (sigma_I-1)*loge(pIindex(t)/pLindex(t))
             + err10 - 1 ;

! eq. 4.52                                                                                  !
equation share_EI_eq  (all,t,alltime)
slack11(t) = - loge(share_EI(t))
             + loge(gamma_EI) + (sigma_I-1)*loge(AI)
             + (sigma_I-1)*loge(pIindex(t)/phatEindex(t))
             + err11 - 1 ;

! eq. 4.53                                                                                  !
equation share_MI_eq  (all,t,alltime)
slack12(t) = - loge(share_MI(t))
             + loge(1-gamma_KI-gamma_LI-gamma_EI)
             + (sigma_I-1)*loge(AI)
             + (sigma_I-1)*loge(pIindex(t)/phatMindex(t))
             + err12 - 1 ;
!-------------------------------------------------------------------------------------------!
! Producer Model: Unit Cost Equations                                                       !
!-------------------------------------------------------------------------------------------!

! eq. 4.47                                                                                  !
```

```
equation unitcost_E  (all,t,alltime)
slack13(t) = -  pEindex(t)
              +  1/AE * ( gamma_KE * (pKindex(t)^(1 - sigma_E))
                        + gamma_LE * (pLindex(t)^(1 - sigma_E))
                        + gamma_EE * (pEindex(t)^(1 - sigma_E))
                        + (1-gamma_KE-gamma_LE-gamma_EE)
                          * (phatMindex(t)^(1 - sigma_E))
                        )^( 1/(1 - sigma_E))
               + err13  - 1  ;

! eq. 4.48                                                                    !
equation unitcost_M  (all,t,alltime)
slack14(t) = - pMindex(t)
              + 1/AM * ( gamma_KM * (pKindex(t)^(1 - sigma_M))
                        + gamma_LM * (pLindex(t)^(1 - sigma_M))
                        + gamma_EM * (phatEindex(t)^(1 - sigma_M))
                        + (1-gamma_KM-gamma_LM-gamma_EM)
                            *(pMindex(t)^(1 - sigma_M))
                         )^( 1/(1 - sigma_M))
              + err14  - 1  ;

! eq. 4.49                                                                    !
equation unitcost_I  (all,t,alltime)
slack15(t) = - pIindex(t)
              + 1 / AI * ( gamma_KI * (pKindex(t)^(1 - sigma_I))
                        + gamma_LI * (pLindex(t)^(1 - sigma_I))
                        + gamma_EI * (phatEindex(t)^(1 - sigma_I))
                        + (1-gamma_KI-gamma_LI-gamma_EI)
                            * (phatMindex(t)^(1 - sigma_I))
                         )^( 1/(1 - sigma_I))
               + err15  - 1  ;
!-------------------------------------------------------------------------------------------!
! Condumer Model                                                                            !
!-------------------------------------------------------------------------------------------!

! eq. 4.3                                                                     !
equation phatE_eq  (all,t,alltime)
slack16(t) = - phatEindex(t)
              + pEindex(t)*( 1 + tax_E(t)) ;

! eq. 4.3                                                                     !
equation phatM_eq  (all,t,alltime)
```

```
slack17(t) = - phatMindex(t)
             + pMindex(t)*( 1 + tax_M(t)) ;

! eq. 4.15                                                              !
equation suby_eq  (all,t,alltime)
slack18(t) = - sup_income(t)
             - income(t) - cbark*pKindex(t) - cbarl*pLindex(t)
              - cbare*phatEindex(t) - cbarm*phatMindex(t) ;


! eq. 4.70                                                              !
equation ddM_eq  (all,t,alltime)
slack19(t) = + alphaM*sup_expend(t)
             - phatMindex(t)*qMCindex(t) + cbarm*phatMindex(t)
             + err16 - 1 ;

! eq. 4.67                                                              !
equation ddK_eq  (all,t,alltime)
slack20(t) = + alphaK*sup_expend(t)
             - pKindex(t)*qKCindex(t) + cbark*pKindex(t)
            + err17 - 1 ;

! eq. 4.68                                                              !
equation ddL_eq  (all,t,alltime)
slack21(t) = + alphaL*sup_expend(t)
             - pLindex(t)*qLCindex(t) + cbarl*pLindex(t)
             + err18 - 1 ;

! eq. 4.69                                                              !
equation ddE_eq  (all,t,alltime)
slack22(t) = + alphaE*sup_expend(t)
             - phatEindex(t)*qECindex(t) + cbare*phatEindex(t)   ;

! eq. 4.33                                                              !
equation income_eq  (all,t,alltime)
slack23(t) = - income(t)
             + (1-tax_L(t))*pLindex(t)*qLindex(t)
             + pKindex(t)*qKindex(t)*adjust_pk*(1 - tax_K(t))
             - dep_rate*qKindex(t)*pIindex(t) + subsidy(t) ;

! eq. 4.32                                                              !
equation expd_eq  (all,t,alltime)
```

```
slack24(t) = - expend(t)
             + pKindex(t)*qKCindex(t) + pLindex(t)*qLCindex(t)
             + phatEindex(t)*qECindex(t) + phatMindex(t)*qMCindex(t) ;

! eq. 4.33                                                                    !
equation savings_eq  (all,t,alltime)
slack25(t) = - savings(t)
             + income(t) - expend(t) + dep_rate*qKindex(t)*pIindex(t)  ;

! eq. 4.37                                                                    !
equation subsidy_eq  (all,t,alltime)
slack26(t) = - subsidy(t)
             + tax_L(t)*pLindex(t)*qLindex(t)
             + tax_K(t)*pKindex(t)*qKindex(t)*adjust_pk
             + tax_E(t)*pEindex(t)*( qECindex(t) + qEMindex(t) + qEIindex(t))
             + tax_M(t)*pMindex(t)*(qMCindex(t)+qMEindex(t) + qMIindex(t));

! eq. 4.34                                                                    !
equation sav_eq  (all,t,alltime)
slack27(t) = - savings(t)
             + pIindex(t)*qIindex(t) ;

! eq. 4.31                                                                    !
equation rental_eq  (all,t,alltime)
slack28(t) = - pKindex(t)*adjust_pk*(1 - tax_K(t))
             + (int_rate(t) + dep_rate)*pIindex(t) ;
!-----------------------------------------------------------------------------------------!
! Market Clearing Consitions                                                  !
!-----------------------------------------------------------------------------------------!

! eq. 4.41                                                                    !
equation mktcls_K  (all,t,alltime)
slack29(t) = - qKindex(t)*adjust_pk
             + qKEindex(t) + qKMindex(t) + qKCindex(t) + qkIindex(t) ;

! eq. 4.42                                                                    !
equation mktcls_L  (all,t,alltime)
slack30(t) = - walras(t)
             -    qLindex(t) + qLEindex(t)
             + qLMindex(t) + qLCindex(t) + qLIindex(t)  ;

! eq. 4.38                                                                    !
```

```
equation mktcls_E  (all,t,alltime)
slack31(t) = - qEindex(t)
             + qEEindex(t) + qEMindex(t) + qECindex(t) + qEIindex(t)  ;

! eq. 4.39                                                                    !
equation mktcls_M  (all,t,alltime)
slack32(t) = - qMindex(t)
             + qMEindex(t) + qMMindex(t) + qMCindex(t) + qMIindex(t)  ;

! eq. 4.40                                                                    !
equation mktcls_I  (all,t,alltime)
slack33(t) = - qGindex(t)
             + qIindex(t) ;
!-------------------------------------------------------------------------------------------!
! Consumer Model: Input Demand Conversion Equations                           !
!-------------------------------------------------------------------------------------------!

! eq. 4.46                                                                    !
equation dd_KE  (all,t,alltime)
slack34(t) = - share_KE(t)
             + pKindex(t)*qKEindex(t) / (pEindex(t)*qEindex(t))   ;

! eq. 4.46                                                                    !
equation dd_LE  (all,t,alltime)
slack35(t) = - share_LE(t)
             + pLindex(t)*qLEindex(t) / (pEindex(t)*qEindex(t))   ;

! eq. 4.46                                                                    !
equation dd_EE  (all,t,alltime)
slack36(t) = - share_EE(t)
             + pEindex(t)*qEEindex(t) / (pEindex(t)*qEindex(t))  ;

! eq. 4.46                                                                    !
equation dd_ME  (all,t,alltime)
slack37(t) = - share_ME(t)
             + phatMindex(t)*qMEindex(t) / (pEindex(t)*qEindex(t))  ;

! eq. 4.46                                                                    !
equation dd_KM  (all,t,alltime)
slack38(t) = - share_KM(t)
             + pKindex(t)*qKMindex(t) / (pMindex(t)*qMindex(t))  ;
```

```
! eq. 4.46                                                                    !
equation dd_LM  (all,t,alltime)
slack39(t) = - share_LM(t)
             + pLindex(t)*qLMindex(t) / (pMindex(t)*qMindex(t))  ;

! eq. 4.46                                                                    !
equation dd_EM  (all,t,alltime)
slack40(t) = - share_EM(t)
             + phatEindex(t)*qEMindex(t) / (pMindex(t)*qMindex(t))   ;

! eq. 4.46                                                                    !
equation dd_MM  (all,t,alltime)
slack41(t) = - share_MM(t)
             + pMindex(t)*qMMindex(t) / (pMindex(t)*qMindex(t))   ;

! eq. 4.46                                                                    !
equation dd_KI  (all,t,alltime)
slack42(t) = - share_KI(t)
             + pKindex(t)*qKIindex(t) / (pIindex(t)*qGindex(t))   ;

! eq. 4.46                                                                    !
equation dd_LI  (all,t,alltime)
slack43(t) = - share_LI(t)
             + pLindex(t)*qLIindex(t) / (pIindex(t)*qGindex(t))   ;

! eq. 4.46                                                                    !
equation dd_EI  (all,t,alltime)
slack44(t) = - share_EI(t)
             + phatEindex(t)*qEIindex(t) / (pIindex(t)*qGindex(t))  ;

! eq. 4.46                                                                    !
equation dd_MI  (all,t,alltime)
slack45(t) = - share_MI(t)
             + phatMindex(t)*qMIindex(t) / (pIindex(t)*qGindex(t)) ;
!-------------------------------------------------------------------------------------------!
! Intra Period Equation                                                                     !
!-------------------------------------------------------------------------------------------!

! eq. 4.26                                                                    !
equation vuindex_eq  (all,t,alltime)
slack46(t) = - sup_expend(t)
             + disc_rate*pvwealth(t)  ;
```

```
!-------------------------------------------------------------------------------------------!
! Intertemporal Equations                                                                   !
!-------------------------------------------------------------------------------------------!
! eq. 4.30                                                                                  !
equation cap_eq  (all,t,fwdtime)
slack47(t) = - qKindex(t+1)
             + dt(t)*qIindex(t) + (1 - dt(t)*dep_rate)*qKindex(t) ;

! eq. 4.26                                                                                  !
equation pvw_eq  (all,t,fwdtime)
slack48(t) = - pvwealth(t)
             + sup_income(t)*(1 - exp(-int_rate(t)*dt(t)))/int_rate(t)
             + pvwealth(t+1)*exp(-int_rate(t)*dt(t)) ;
!-------------------------------------------------------------------------------------------!
! Transversality Conditions                                                                 !
!-------------------------------------------------------------------------------------------!

! eq. 4.44                                                                                  !
equation sstate1_eq  (all,t,begintime)
slack49(t) = - qKindex(t)
             + qKindex0 ;

! eq. 4.43                                                                                  !
equation sstate2_eq  (all,t,endtime)
slack50(t) = - qIindex(t)
             + dep_rate*qKindex(t) ;
!-------------------------------------------------------------------------------------------!
! Components of Equivalent Variation Expression                                             !
!-------------------------------------------------------------------------------------------!

! eq. 4.25                                                                                  !
equation sv_eq  (all,t,alltime)
slack51(t)  = - sv(t)
              + cbark*pKindex(t) + cbarl*pLindex(t)
              + cbare*phatEindex(t) + cbarm*phatMindex(t) ;

! eq. 4.25                                                                                  !
equation omega_eq  (all,t,fwdtime)
slack52(t) = - omega(t)
             + sv(t)*(1 - exp(-int_rate(t)*dt(t)))/int_rate(t)
             + omega(t+1)*exp(-int_rate(t)*dt(t)) ;
```

```
! eq. 4.25                                                                    !
equation omend_eq  (all,t,endtime)
slack53(t) = - omega(t)
             + sv(t)/int_rate(t) ;

! eq. 5.1                                                                     !
equation pvwhat_eq  (all,t,alltime)
slack54(t) = - pvwhat(t)
             + pvwealth(t) - omega(t) ;

! eq. 5.3                                                                     !
equation lw_eq  (all,t,alltime)
slack55(t) = - lw(t)
             + loge(disc_rate*pvwhat(t)/phat(t)) ;

! eq. 4.18                                                                    !
equation phat_eq  (all,t,alltime)
slack56(t) = - phat(t)
             + ((pKindex(t)/alphaK)^alphaK)
               *((pLindex(t)/alphaL)^alphaL)
               *((phatEindex(t)/alphaE)^alphaE)
               *((phatMindex(t)/alphaM)^alphaM) ;

! eq. 5.5                                                                     !
equation V_eq  (all,t,fwdtime)
slack57(t) = - V(t)
             + lw(t)*(1 - exp(-disc_rate*dt(t)))/disc_rate
             + V(t+1)*exp(-disc_rate*dt(t)) ;

! eq. 5.5                                                                     !
equation Vend_eq  (all,t,endtime)
slack58(t) = - V(t)
             + lw(t)/disc_rate ;

! eq. 5.5                                                                     !
equation theta_eq  (all,t,fwdtime)
slack59(t) = - theta(t)
             + loge(disc_rate) - disc_rate*dt(t) + int_rate(t)*dt(t) - loge(phat(t)) ;

! eq. 5.6                                                                     !
equation the_eq  (all,t,endtime)
slack60(t) = - theta(t)
```

```
                + loge(disc_rate) - loge(phat(t)) ;

! eq. 5.6                                                                    !
equation phi_eq  (all,t,fwdtime)
slack61(t) = - phi(t)
               + theta(t)*(1 - exp(-disc_rate*dt(t)))/disc_rate
               + phi(t+1)*exp(-disc_rate*dt(t)) ;

! eq. 5.6                                                                    !
equation phiend_eq  (all,t,endtime)
slack62(t) = - phi(t)
               + theta(t)/disc_rate ;

! eq. 4.13                                                                   !
equation util_eq  (all,t,alltime)
slack63(t) = - utility(t)
               + ((qKCindex(t)  - cbark)^alphak)
                  *((qLCindex(t)  - cbarl)^alphal)
                  *((qECindex(t)  - cbare)^alphae)
                  *((qMCindex(t)  - cbarm)^alpham) ;
!-------------------------------------------------------------------------------------------!
! End of TABLO Program                                                                      !
! Calculation of Equivalent Variation requires before and after policy reform               !
! models. The above equations represents for only one regime.                               !
!-------------------------------------------------------------------------------------------!
```

Bibliography

Arndt, Channing, (1996): "An Introduction to Systematic Sensitivity Analysis via Gaussian Quadrature," GTAP Technical Paper No.2, Global Trade Analysis Project, Purdue University, July.

Arndt, Channing, and K. R. Pearson (1998): "How to Carry out Systematic Sensitivity Analysis via Gaussian Quadrature and GEMPACK," GTAP Technical Paper No.3, Global Trade Analysis Project, Purdue University, April.

Ballard, Charles L., and Steven G. Medema (1993): "The Marginal Efficiency Effects of Taxes and Subsidies in the Presence of Externalities," *Journal of Public Economics*, 52, pp. 199-216.

Bernheim, Douglas, John K. Scholz, and John B. Shoven: "Consumption Taxation in a General Equilibrium Model: How reliable are simulation results?"

Bovenberg, A.Lans. and Ruud A. de Mooij (1994): "Environmental Levies and Distortionary Taxation," *American Economic Review*, 94, September, pp.1085-89.

Bovenberg, A.Lans. and Lawrence.H. Goulder (1994): "Optimal Environmental Taxation in the Presence of Other Taxes: General Equilibrium Analyses," Working Paper No. 4897, National Bureau of Economic Research, Inc.

Chiang, Alpha C. (1992): "Elements of Dynamic Optimization," McGraw-Hill, Inc.

Codi, G, and K. R. Pearson, and Peter J. Wilcoxen (1992): "General-Purpose Software for Intertemporal Economics Models," *Computer Science in Economics and Management*, 5, pp.57-79.

Codi, G, and K. R. Pearson (1988): "GEMPACK: General-Purpose Software for Applied General Equilibrium and Other Economic Modellers," *Computer Science in Economics and Management*, 1, pp.189-207.

Deaton, Angus, and John Muellbauer (1982): "Economics and Consumer Behavior," Cambridge University Press, Cambridge.

Eckstein, Otto (1983): "The DRI Model of the US Economy," McGraw-Hill, Inc.

Fishman, George S. (1996): "Monte Carlo: Concepts, Algorithms, and Applications," Springer-Verlag, Springer Series in Operations Research, New York

Fossatoi, Amedeo (ed.) (1996): "Economic Modelling Under the Applied General Equilibrium Approach," Avebury, England.

Fullerton, Don, and Gilbert E. Metcalf (1998): "Environmental Controls, Scarcity Rents, and Preexisting Distortions," *Journal of Public Economics*, 80(2), pp. 249-67.

Fullerton, Don, and Gilbert E. Metcalf (1998): "Environmental Taxes and The Double Dividend Hypothesis: Did You Really Expect Something for Nothing?" *Chicago Kent Law Review*, 79(1), pp. 221-56.

Fullerton, Don (199): "Environmental Levies and Distortionary Taxation: Comment," *American Economic Review*, 87, pp. 245-51.

GAUSS, "The GAUSS System Version 3.0", Aptech Systems, Inc., Washington

Goulder, Lawrence H. (1995): "Environmental Taxation and the 'Double Dividend': A Reader's Guide," Working Paper No. 4896, National Bureau of Economic Research, Inc.

GPD-1, (1994): "An Introduction to GEMPACK," second edition, Centre of Policy Studies and Impact Project, Monash University, April,

GPD-2, (1994): "User's Guide to TABLE, GEMSIM, and TABLO-Generated Programs", second edition, Centre of Policy Studies and Impact Project, Monash University, April,

Hall, Bronwyn H., and Schnake Rebecca (1986): "TSP User's Manual," version 4.0, TSP International

Harrison, Glenn W., Richard Jones, Larry J. Kimbell, and Randall Wigle (1993): "How Robust is Applied General Equilibrium Analysis," *Journal of Policy Modeling*, 15(1), pp. 99-115.

Harrison, Glenn W. and H. D. Vinod (1992): "The Sensitivity Analysis of Applied General Equilibrium Model: Completely Randomized Factorial Sampling Designs," *The Review of Economics and Statistics*, 74(2), May, pp. 357-62.

Horridge, J. M., and B. R. Parmenter, and K. R. Pearson (1993): "ORANI-F: A General Equilibrium Model of the Australian Economy," *Economic and Financial Computing*, 3, pp.71-140.

Johnston, John (1984): "Econometric Methods," McGraw-Hill, New York

Jorgenson, Dale W. (1984): "Econometric Methods for Applied General Equilibrium Analysis," Applied General Equilibrium Analysis, H. Scarf and J. Shoven (eds.), Cambridge University Press, pp. 139-207

Jorgenson, Dale W., and Peter J. Wilcoxen (1993): "Reducing US Carbon Emissions: An Econometric General Equilibrium Assessment," *Resources and Energy Economics*, 15(1), pp. 7-26.

Jorgenson, Dale W., and Peter J. Wilcoxen (1990): "Environmental Regulation and U.S. Economic Growth," *RAND Journal of Economics*, 21(2), pp. 314-40.

Jorgenson, Dale W., and Zvi Griliches (1971): "Divisia Index Numbers and Productivity Measurement," *Review of Income and Wealth*, 17(2) pp. 227-29

Jorgenson, Dale W., Daniel, T. Slesnick, and Peter J. Wilcoxen (1992): "Carbon Taxes and Economic Welfare," Brookings Papers: Microeconomics

Kendrick, David A. (1988): "Feedback: A New Framework for Macroeconomic Policy," Kluwer Academic Publishers

Kendrick, David A. (1981): "Stochastic Control for Economic Models," McGraw-Hill Book Company, New York

Kim, Seung-Rae (2003): "Essays on Interactions Between Environmental and Fiscal Policies: Analytical and Numerical General Equilibrium Analyses," Ph.D. Dissertation, University of Texas at Austin, Austin

Lee, Dwight R. and Walter S. Misiolek (1986): "Substituting Pollution Taxation for General Taxation: Some Implications for Efficiency in Pollution Taxation," *Journal of Environmental Economics and Management*, November, 29(3), pp. S64-S77.

Lluch, Constantino, Alan A. Powell, Ross A. Williams (1977): "Patterns in Household Demand and Saving," Oxford University Press, New York

Lucas, Robert Jr. (1976): "Economic Policy Evaluation: A Critique," in The Phillips Curve and Labor Markets, K. Brunner and A. H. Meltzer (eds.), North-Holland Publishing Co., Amsterdam, pp. 525-46.

Mansur, Ahsan, and John Whalley (19): "Numerical Specification of Applied General Equilibrium Models: estimation, calibration, and data,"

McKibbin, Warwick J., and Peter J. Wilcoxen (1997): "A Better way to slow Global Change,"

McKitrick, Ross R. (1995): "The Econometric Critique of Computable General Equilibrium Modeling: The Role of Parameter Estimation," Discussion Paper No.: 95-27, Department of Economics, The University of British Columbia, Vancouver, Canada.

Morgam, M. Granger, and Max Henrion, "Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis," Cambridge University Press, Cambridge

Oates Wallace E. (1995): "Green Taxes: Can We Protect the Environment and Improve the Tax System at the Same Time," *Southern Economic Journal*, April, 61(4), pp. 915-22

O'Neill R. V., R. H. Gardner, and J. B. Mankin (1980): "Analysis of Parameter Error in a Nonlinear Model," *Ecological Modelling*, 8, pp. 297-311.

Pagan, Adrian R., and John H. Shannon (1985): "Sensitivity Analysis for Linearized Computable General Equilibrium Models," in John Piggot and John Whalley (eds.), New Developments in Applied General Equilibrium Analysis, Cambridge University Press, pp. 104-18

Parry, Ian W. H. (1995): "Pollution Taxes and Revenue Recycling," *Journal of Environmental Economics and Management*, November, 29(3), pp. S64-S77.

Pearson, K. R. (1992): "Simulation Software for Use with Notes and Problems in Applied General Equilibrium Economics," North-Holland, Amsterdam.

Tuladhar, Sugandha D., and Peter J. Wilcoxen (1998): "An Econometric Look at the Double Dividend Hypothesis," *National Tax Association Proceedings*, Washington DC, 1999.

Schwartz, Randal L. (1993): "Learning Perl," 1st Edition, O'Reilly and Associates.

Smith, B. G. (1993) “Optimal Regional Economics Policies Using Computable General Equilibrium Models,” Ph.D. Dissertation, Department of Economics, The University of Texas, Austin, Texas.

Varian, Hall R. (1992): "*Microeconomic Analysis*," Third Edition, W. W. Norton & Company, Inc., New York.

Whalley, John (1984): "Hidden Challenges in Recent Applied General Equilibrium Exercises," in John Piggot and John Whalley (eds.), *New Developments in Applied General Equilibrium Analysis*, Cambridge University Press, pp. 25-41

Webster Mort., Andrei Sokolov (1998): "Quantifying the Uncertainty in Climate Predictions," MIT Global Change Joint Program Report No. 37, Massachusetts Institute of Technology, Joint Program on the Science and Policy of Global Change.

Wlcoxen, Peter J. (1989): "The Effects of Environmental Regulation and Energy Prices on U.S. Economic Performance," Ph.D. Dissertation, Harvard University.

Wilcoxen Peter J. (1995): "Intertemporal Optimization in General Equilibrium: A Practical Introduction," Department of Economics, The University of Texas at Austin, August.

Vita

Sugandha Dhar Tuladhar was born on December 25, 1963 in Kathmandu, Nepal, the son of Tirtha Dhar Tuladhar and late Pun Tara Tuladhar. He received a Bachelor of Science in Mechanical Engineering from Regional Engineering College, Rourkela, India, in June 1988. Following graduation, he was employed as a researcher by Economic Services Centre, Kathmandu, Nepal till 1992.

In August 1992, he entered the Operations Research and Industrial Engineering Program at the University of Texas at Austin as a Fulbright Scholar. He received Master of Science in May 1994. Subsequently, he entered the Doctoral program at the Department of Economics, University of Texas, in August 1994 and earned his Master's degree Economics in 1996. Prior to completing his dissertation, *Confidence Interval for Computable General Equilibrium Models*, in 2003, he worked as a financial sector economist in Nepal for three years for the Asian Development Bank.

Permanent address: 9/328 Bhedasingh, Kathmandu, Nepal

This dissertation was typed by the author.